

Please type a plus sign (+) inside this box



06-29-00

Box 560 A
PTO/SB/05 (2/98)

06/27/00

03605793-062700
06/27/00UTILITY
PATENT APPLICATION
TRANSMITTAL

(Only for nonprovisional applications under 37 CFR § 1.53(b))

Attorney Docket No.

210121.427C16

First Inventor or Application Identifier

Jiangchun Xu

Title

COMPOSITIONS AND METHODS FOR THE THERAPY
AND DIAGNOSIS OF PROSTATE CANCER

Express Mail Label No.

EL615231109US

APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

ADDRESS TO:

Box Patent Application
Assistant Commissioner for Patents
Washington, D.C. 202311. ☐ General Authorization Form & Fee Transmittal
(Submit an original and a duplicate for fee processing)2. ☒ Specification [Total Pages] **206**
(preferred arrangement set forth below)
- Descriptive Title of the Invention
- Cross References to Related Applications
- Statement Regarding Fed sponsored R & D
- Reference to Microfiche Appendix
- Background of the Invention- Brief Summary of the Invention
- Brief Description of the Drawings (if filed)
- Detailed Description
- Claim(s)
- Abstract of the Disclosure3. ☒ Drawing(s) (35 USC 113) [Total Sheets] **16**4. Oath or Declaration [Total Pages] ☐
a. ☐ Newly executed (original or copy)
b. ☐ Copy from a prior application (37 CFR 1.63(d))
(for continuation/divisional with Box 17 completed)i. ☐ DELETION OF INVENTOR(S)
Signed statement attached deleting
inventor(s) named in the prior application,
see 37 CFR 1.63(d)(2) and 1.33(b)5. ☐ Incorporation By Reference (useable if box 4b is
checked) The entire disclosure of the prior application,
from which a copy of the oath or declaration is supplied
under Box 4b, is considered to be part of the disclosure of
the accompanying application and is hereby incorporated
by reference therein.6. ☐ Microfiche Computer Program (Appendix)
7. Nucleotide and Amino Acid Sequence Submission
(if applicable, all necessary)a. ☒ Computer-Readable Copy
b. ☒ Paper Copy (identical to computer copy)
c. ☒ Statement verifying identity of above copies

ACCOMPANYING APPLICATION PARTS

- 8.
- ☐
- Assignment Papers (cover sheet & document(s))
-
- 9.
- ☐
- 37 CFR 3.73(b) Statement (when there is an assignee)
- ☐
- Power of Attorney
-
- 10.
- ☐
- English Translation Document (if applicable)
-
- 11.
- ☐
- Information Disclosure Statement (IDS)/PTO-1449
- ☐
- Copies of IDS Citations
-
- 12.
- ☐
- Preliminary Amendment
-
- 13.
- ☒
- Return Receipt Postcard
-
- 14.
- ☐
- Small Entity Statement(s)
- ☐
- Statement filed in prior application, Status still proper and desired
-
- 15.
- ☐
- Certified Copy of Priority Document(s) (if foreign priority is claimed)
-
- 16.
- ☒
- Other:
- Certificate of Express Mail

17. If a CONTINUING APPLICATION, check appropriate box and supply the requisite information below and in a preliminary amendment

☐ Continuation ☐ Divisional ☒ Continuation-In-Part (CIP) of prior Application No.: **09/593,793**Prior application information: Examiner not assigned Group / Art Unit not assigned☐ Claims the benefit of Provisional Application No. _____

CORRESPONDENCE ADDRESS

Jane E. R. Potter
Seed Intellectual Property Law Group PLLC
701 Fifth Avenue, Suite 6300
Seattle, Washington 98104-7092
Phone: (206) 622-4900 Fax: (206) 682-6031

Respectfully submitted,

TYPED or PRINTED NAME Jane E. R. PotterSIGNATURE Jane E. R. PotterREGISTRATION NO. **33,332**Date June 27, 2000

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

PATENT

Applicants : Jiangchun Xu, Bellevue, WA; Davin C. Dillon, Redmond, WA; Jennifer L. Mitcham, Redmond, WA; Susan L. Harlocker, Seattle, WA; Yuqiu Jiang, Kent, WA; Michael D. Kalos, Seattle, WA; Gary R. Fanger, Mill Creek, WA; Marc W. Retter, Carnation, WA; John A. Stolk, Bothell, WA; Craig H. Day, Seattle, WA; Thomas S. Vedvick, Bainbridge Island, WA; Darrick Carter, Seattle, WA; Samuel Li, Redmond, WA; Aijun Wang, Issaquah, WA; Yasir A. W. Skeiky, Seattle, WA; William Hepler, Seattle, WA; Robert A. Henderson, Seattle, WA

Filed : June 27, 2000

For : COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF PROSTATE CANCER

Docket No. : 210121.427C16

Date : June 27, 2000

Box Patent Application
Assistant Commissioner for Patents
Washington, DC 20231

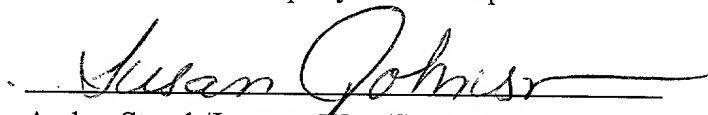
CERTIFICATE OF MAILING BY "EXPRESS MAIL"

Assistant Commissioner for Patents:

I hereby certify that the enclosures listed below are being deposited with the United States Postal Service "EXPRESS MAIL Post Office to Addressee" service under 37 C.F.R. § 1.10, Mailing Label Certificate No. EL615231109US, on June 27, 2000, addressed to Box Patent Application, Assistant Commissioner for Patents, Washington, DC 20231.

Respectfully submitted,

Seed Intellectual Property Law Group PLLC


Amber Straub/Jeanette West/Susan Johnson

Enclosures:

Postcard
Form PTO/SB/05
Specification, Claims, Abstract (206 pages)
16 Sheets of Drawings (Figures 1-12)
Sequence Listing (369 pages)
Declaration for Sequence Listing
Diskette for Sequence Listing

u:\sharons\210121\427



COMPOSITIONS AND METHODS FOR THE THERAPY
AND DIAGNOSIS OF PROSTATE CANCER

CROSS REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of U.S. Patent Application
5 No. 09/593,793, filed June 13, 2000, which is a continuation-in-part of U.S. Patent
Application No. 09/_____, filed May 12, 2000, which is a continuation-in-part of U.S.
Patent Application No. 09/568,100, filed May 9, 2000, which is a continuation-in-part of
U.S. Patent Application No. 09/536,857, filed March 27, 2000, which is a continuation-in-
part of U. S. Patent Application No. 09/483,672, filed January 14, 2000, which is a
10 continuation-in-part of U.S. Patent Application No. 09/439,313, filed November 12, 1999,
which is a continuation-in-part of U.S. Patent Application No. 09/352,616, filed July 13,
1999, which is a continuation-in-part of U.S. Patent Application No. 09/288,946, filed
April 9, 1999, which is a continuation-in-part of U.S. Patent Application No. 09/232,149,
filed January 15, 1999, which is a continuation-in-part of U.S. Patent Application No.
15 09/159,812, filed September 23, 1998, which is a continuation-in-part of U.S. Patent
Application No. 09/115,453, filed July 14, 1998, which is a continuation-in-part of U.S.
Patent Application No. 09/030,607, filed February 25, 1998, which is a continuation-in-part
of U.S. Patent Application No. 09/020,956, filed February 9, 1998, which is a continuation-
in-part of U.S. Patent Application No. 08/904,804, filed August 1, 1997, which is a
20 continuation-in-part of U.S. Patent Application No. 08/806,099, filed February 25, 1997.

TECHNICAL FIELD OF THE INVENTION

The present invention relates generally to therapy and diagnosis of cancer,
such as prostate cancer. The invention is more specifically related to polypeptides
comprising at least a portion of a prostate-specific protein, and to polynucleotides encoding
25 such polypeptides. Such polypeptides and polynucleotides may be used in compositions
for prevention and treatment of prostate cancer, and for the diagnosis and monitoring of
such cancers.

BACKGROUND OF THE INVENTION

Cancer is a significant health problem throughout the world. Although Cancer is a significant health problem throughout the world. Although advances have been made in detection and therapy of cancer, no vaccine or other universally successful method
5 for prevention or treatment is currently available. Current therapies, which are generally based on a combination of chemotherapy or surgery and radiation, continue to prove inadequate in many patients.

Prostate cancer is the most common form of cancer among males, with an estimated incidence of 30% in men over the age of 50. Overwhelming clinical evidence
10 shows that human prostate cancer has the propensity to metastasize to bone, and the disease appears to progress inevitably from androgen dependent to androgen refractory status, leading to increased patient mortality. This prevalent disease is currently the second leading cause of cancer death among men in the U.S.

In spite of considerable research into therapies for the disease, prostate
15 cancer remains difficult to treat. Commonly, treatment is based on surgery and/or radiation therapy, but these methods are ineffective in a significant percentage of cases. Two previously identified prostate specific proteins - prostate specific antigen (PSA) and prostatic acid phosphatase (PAP) - have limited therapeutic and diagnostic potential. For example, PSA levels do not always correlate well with the presence of prostate cancer,
20 being positive in a percentage of non-prostate cancer cases, including benign prostatic hyperplasia (BPH). Furthermore, PSA measurements correlate with prostate volume, and do not indicate the level of metastasis.

In spite of considerable research into therapies for these and other cancers, prostate cancer remains difficult to diagnose and treat effectively. Accordingly, there is a
25 need in the art for improved methods for detecting and treating such cancers. The present invention fulfills these needs and further provides other related advantages.

SUMMARY OF THE INVENTION

Briefly stated, the present invention provides compositions and methods for the diagnosis and therapy of cancer, such as prostate cancer. In one aspect, the present invention provides polypeptides comprising at least a portion of a prostate-specific protein, or a variant thereof. Certain portions and other variants are immunogenic, such that the ability of the variant to react with antigen-specific antisera is not substantially diminished. Within certain embodiments, the polypeptide comprises a sequence that is encoded by a polynucleotide sequence selected from the group consisting of: (a) sequences recited in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824; (b) variants of a sequence recited in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824; and (c) complements of a sequence of (a) or (b). In specific embodiments, the polypeptides of the present invention comprise at least a portion of a tumor protein that includes an amino acid sequence selected from the group consisting of sequences recited in SEQ ID NO: 112-114, 172, 176, 178, 327, 329, 331, 336, 339, 376-380, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534, 537-551, 553-568, 573-586, 588-590, 592, 706-708, 775, 776, 778, 780, 781, 811, 814, 818, 826 and 827, and variants thereof.

The present invention further provides polynucleotides that encode a polypeptide as described above, or a portion thereof (such as a portion encoding at least 15 amino acid residues of a prostate-specific protein), expression vectors comprising such polynucleotides and host cells transformed or transfected with such expression vectors.

Within other aspects, the present invention provides pharmaceutical compositions comprising a polypeptide or polynucleotide as described above and a physiologically acceptable carrier.

Within a related aspect of the present invention, immunogenic compositions, or vaccines for prophylactic or therapeutic use are provided. Such

compositions comprise a polypeptide or polynucleotide as described above and an immunostimulant.

The present invention further provides pharmaceutical compositions that comprise: (a) an antibody or antigen-binding fragment thereof that specifically binds to a prostate-specific protein; and (b) a physiologically acceptable carrier.

Within further aspects, the present invention provides pharmaceutical compositions comprising: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) a pharmaceutically acceptable carrier or excipient. Antigen presenting cells include dendritic cells, macrophages, monocytes, fibroblasts and B cells.

Within related aspects, immunogenic compositions, or vaccines, are provided that comprise: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) an immunostimulant.

The present invention further provides, in other aspects, fusion proteins that comprise at least one polypeptide as described above, as well as polynucleotides encoding such fusion proteins.

Within related aspects, pharmaceutical compositions comprising a fusion protein, or a polynucleotide encoding a fusion protein, in combination with a physiologically acceptable carrier are provided.

Compositions are further provided, within other aspects, that comprise a fusion protein, or a polynucleotide encoding a fusion protein, in combination with an immunostimulant.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient a composition as recited above. The patient may be afflicted with prostate cancer, in which case the methods provide treatment for the disease, or patient considered at risk for such a disease may be treated prophylactically.

The present invention further provides, within other aspects, methods for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a prostate-specific protein, wherein the step of

contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the protein from the sample.

Within related aspects, methods are provided for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated as
5 described above.

Methods are further provided, within other aspects, for stimulating and/or expanding T cells specific for a prostate-specific protein, comprising contacting T cells with one or more of: (i) a polypeptide as described above; (ii) a polynucleotide encoding such a polypeptide; and/or (iii) an antigen presenting cell that expresses such a polypeptide;
10 under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells. Isolated T cell populations comprising T cells prepared as described above are also provided.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective
15 amount of a T cell population as described above.

The present invention further provides methods for inhibiting the development of a cancer in a patient, comprising the steps of: (a) incubating CD4⁺ and/or CD8⁺ T cells isolated from a patient with one or more of: (i) a polypeptide comprising at least an immunogenic portion of a prostate-specific protein; (ii) a polynucleotide encoding
20 such a polypeptide; and (iii) an antigen-presenting cell that expressed such a polypeptide; and (b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of a cancer in the patient. Proliferated cells may, but need not, be cloned prior to administration to the patient.

Within further aspects, the present invention provides methods for
25 determining the presence or absence of a cancer in a patient, comprising: (a) contacting a biological sample obtained from a patient with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; and (c) comparing the amount of polypeptide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

Within preferred embodiments, the binding agent is an antibody, more preferably a monoclonal antibody. The cancer may be prostate cancer.

The present invention also provides, within other aspects, methods for monitoring the progression of a cancer in a patient. Such methods comprise the steps of:

- 5 (a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polypeptide detected in step (c) with the amount detected in step
- 10 (b) and therefrom monitoring the progression of the cancer in the patient.

- The present invention further provides, within other aspects, methods for determining the presence or absence of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a prostate-specific protein; (b) detecting in the
- 15 sample a level of a polynucleotide, preferably mRNA, that hybridizes to the oligonucleotide; and (c) comparing the level of polynucleotide that hybridizes to the oligonucleotide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within certain embodiments, the amount of mRNA is detected via polymerase chain reaction using, for example, at least one
 - 20 oligonucleotide primer that hybridizes to a polynucleotide encoding a polypeptide as recited above, or a complement of such a polynucleotide. Within other embodiments, the amount of mRNA is detected using a hybridization technique, employing an oligonucleotide probe that hybridizes to a polynucleotide that encodes a polypeptide as recited above, or a complement of such a polynucleotide.

- 25 In related aspects, methods are provided for monitoring the progression of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a prostate-specific protein; (b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; (c) repeating steps (a) and (b) using a biological sample

obtained from the patient at a subsequent point in time; and (d) comparing the amount of polynucleotide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

Within further aspects, the present invention provides antibodies, such as
 5 monoclonal antibodies, that bind to a polypeptide as described above, as well as diagnostic kits comprising such antibodies. Diagnostic kits comprising one or more oligonucleotide probes or primers as described above are also provided.

These and other aspects of the present invention will become apparent upon
 reference to the following detailed description and attached drawings. All references
 10 disclosed herein are hereby incorporated by reference in their entirety as if each was incorporated individually.

BRIEF DESCRIPTION OF THE DRAWINGS AND SEQUENCE IDENTIFIERS

Figure 1 illustrates the ability of T cells to kill fibroblasts expressing the
 representative prostate-specific polypeptide P502S, as compared to control fibroblasts. The
 15 percentage lysis is shown as a series of effector:target ratios, as indicated.

Figures 2A and 2B illustrate the ability of T cells to recognize cells
 expressing the representative prostate-specific polypeptide P502S. In each case, the
 number of γ -interferon spots is shown for different numbers of responders. In Figure 2A,
 data is presented for fibroblasts pulsed with the P2S-12 peptide, as compared to fibroblasts
 20 pulsed with a control E75 peptide. In Figure 2B, data is presented for fibroblasts
 expressing P502S, as compared to fibroblasts expressing HER-2/*neu*.

Figure 3 represents a peptide competition binding assay showing that the
 P1S#10 peptide, derived from P501S, binds HLA-A2. Peptide P1S#10 inhibits HLA-A2
 restricted presentation of fluM58 peptide to CTL clone D150M58 in TNF release bioassay.
 25 D150M58 CTL is specific for the HLA-A2 binding influenza matrix peptide fluM58.

Figure 4 illustrates the ability of T cell lines generated from P1S#10
 immunized mice to specifically lyse P1S#10-pulsed Jurkat A2Kb targets and P501S-

transduced Jurkat A2Kb targets, as compared to EGFP-transduced Jurkat A2Kb. The percent lysis is shown as a series of effector to target ratios, as indicated.

Figure 5 illustrates the ability of a T cell clone to recognize and specifically lyse Jurkat A2Kb cells expressing the representative prostate-specific polypeptide P501S, thereby demonstrating that the P1S#10 peptide may be a naturally processed epitope of the P501S polypeptide.

Figures 6A and 6B are graphs illustrating the specificity of a CD8⁺ cell line (3A-1) for a representative prostate-specific antigen (P501S). Figure 6A shows the results of a ⁵¹Cr release assay. The percent specific lysis is shown as a series of effector:target ratios, as indicated. Figure 6B shows the production of interferon-gamma by 3A-1 cells stimulated with autologous B-LCL transduced with P501S, at varying effector:target ratios as indicated.

Figure 7 is a Western blot showing the expression of P501S in baculovirus.

Figure 8 illustrates the results of epitope mapping studies on P501S.

Figure 9 is a schematic representation of the P501S protein showing the location of transmembrane domains and predicted intracellular and extracellular domains.

Figure 10 is a genomic map showing the location of the prostate genes P775P, P704P, B305D, P712P and P774P within the Cat Eye Syndrome region of chromosome 22q11.2

Figure 11 shows the results of an ELISA assay to determine the specificity of rabbit polyclonal antisera raised against P501S.

Figures 12A(1), 12A(2), 12A(3), and B are the full-length cDNA (SEQ ID NO:591) and predicted amino acid (SEQ ID NO:592) sequences, respectively, for the clone P788P.

SEQ ID NO: 1 is the determined cDNA sequence for F1-13
 SEQ ID NO: 2 is the determined 3' cDNA sequence for F1-12
 SEQ ID NO: 3 is the determined 5' cDNA sequence for F1-12
 SEQ ID NO: 4 is the determined 3' cDNA sequence for F1-16
 SEQ ID NO: 5 is the determined 3' cDNA sequence for H1-1

SEQ ID NO: 6 is the determined 3' cDNA sequence for H1-9
 SEQ ID NO: 7 is the determined 3' cDNA sequence for H1-4
 SEQ ID NO: 8 is the determined 3' cDNA sequence for J1-17
 SEQ ID NO: 9 is the determined 5' cDNA sequence for J1-17
 5 SEQ ID NO: 10 is the determined 3' cDNA sequence for L1-12
 SEQ ID NO: 11 is the determined 5' cDNA sequence for L1-12
 SEQ ID NO: 12 is the determined 3' cDNA sequence for N1-1862
 SEQ ID NO: 13 is the determined 5' cDNA sequence for N1-1862
 SEQ ID NO: 14 is the determined 3' cDNA sequence for J1-13
 10 SEQ ID NO: 15 is the determined 5' cDNA sequence for J1-13
 SEQ ID NO: 16 is the determined 3' cDNA sequence for J1-19
 SEQ ID NO: 17 is the determined 5' cDNA sequence for J1-19
 SEQ ID NO: 18 is the determined 3' cDNA sequence for J1-25
 SEQ ID NO: 19 is the determined 5' cDNA sequence for J1-25
 15 SEQ ID NO: 20 is the determined 5' cDNA sequence for J1-24
 SEQ ID NO: 21 is the determined 3' cDNA sequence for J1-24
 SEQ ID NO: 22 is the determined 5' cDNA sequence for K1-58
 SEQ ID NO: 23 is the determined 3' cDNA sequence for K1-58
 SEQ ID NO: 24 is the determined 5' cDNA sequence for K1-63
 20 SEQ ID NO: 25 is the determined 3' cDNA sequence for K1-63
 SEQ ID NO: 26 is the determined 5' cDNA sequence for L1-4
 SEQ ID NO: 27 is the determined 3' cDNA sequence for L1-4
 SEQ ID NO: 28 is the determined 5' cDNA sequence for L1-14
 SEQ ID NO: 29 is the determined 3' cDNA sequence for L1-14
 25 SEQ ID NO: 30 is the determined 3' cDNA sequence for J1-12
 SEQ ID NO: 31 is the determined 3' cDNA sequence for J1-16
 SEQ ID NO: 32 is the determined 3' cDNA sequence for J1-21
 SEQ ID NO: 33 is the determined 3' cDNA sequence for K1-48
 SEQ ID NO: 34 is the determined 3' cDNA sequence for K1-55

SEQ ID NO: 35 is the determined 3' cDNA sequence for L1-2
 SEQ ID NO: 36 is the determined 3' cDNA sequence for L1-6
 SEQ ID NO: 37 is the determined 3' cDNA sequence for N1-1858
 SEQ ID NO: 38 is the determined 3' cDNA sequence for N1-1860
 5 SEQ ID NO: 39 is the determined 3' cDNA sequence for N1-1861
 SEQ ID NO: 40 is the determined 3' cDNA sequence for N1-1864
 SEQ ID NO: 41 is the determined cDNA sequence for P5
 SEQ ID NO: 42 is the determined cDNA sequence for P8
 SEQ ID NO: 43 is the determined cDNA sequence for P9
 10 SEQ ID NO: 44 is the determined cDNA sequence for P18
 SEQ ID NO: 45 is the determined cDNA sequence for P20
 SEQ ID NO: 46 is the determined cDNA sequence for P29
 SEQ ID NO: 47 is the determined cDNA sequence for P30
 SEQ ID NO: 48 is the determined cDNA sequence for P34
 15 SEQ ID NO: 49 is the determined cDNA sequence for P36
 SEQ ID NO: 50 is the determined cDNA sequence for P38
 SEQ ID NO: 51 is the determined cDNA sequence for P39
 SEQ ID NO: 52 is the determined cDNA sequence for P42
 SEQ ID NO: 53 is the determined cDNA sequence for P47
 20 SEQ ID NO: 54 is the determined cDNA sequence for P49
 SEQ ID NO: 55 is the determined cDNA sequence for P50
 SEQ ID NO: 56 is the determined cDNA sequence for P53
 SEQ ID NO: 57 is the determined cDNA sequence for P55
 SEQ ID NO: 58 is the determined cDNA sequence for P60
 25 SEQ ID NO: 59 is the determined cDNA sequence for P64
 SEQ ID NO: 60 is the determined cDNA sequence for P65
 SEQ ID NO: 61 is the determined cDNA sequence for P73
 SEQ ID NO: 62 is the determined cDNA sequence for P75
 SEQ ID NO: 63 is the determined cDNA sequence for P76

SEQ ID NO: 64 is the determined cDNA sequence for P79
 SEQ ID NO: 65 is the determined cDNA sequence for P84
 SEQ ID NO: 66 is the determined cDNA sequence for P68
 SEQ ID NO: 67 is the determined cDNA sequence for P80 (also referred to
 5 as P704P)

SEQ ID NO: 68 is the determined cDNA sequence for P82
 SEQ ID NO: 69 is the determined cDNA sequence for U1-3064
 SEQ ID NO: 70 is the determined cDNA sequence for U1-3065
 SEQ ID NO: 71 is the determined cDNA sequence for V1-3692
 10 SEQ ID NO: 72 is the determined cDNA sequence for 1A-3905
 SEQ ID NO: 73 is the determined cDNA sequence for V1-3686
 SEQ ID NO: 74 is the determined cDNA sequence for R1-2330
 SEQ ID NO: 75 is the determined cDNA sequence for 1B-3976
 SEQ ID NO: 76 is the determined cDNA sequence for V1-3679
 15 SEQ ID NO: 77 is the determined cDNA sequence for 1G-4736
 SEQ ID NO: 78 is the determined cDNA sequence for 1G-4738
 SEQ ID NO: 79 is the determined cDNA sequence for 1G-4741
 SEQ ID NO: 80 is the determined cDNA sequence for 1G-4744
 SEQ ID NO: 81 is the determined cDNA sequence for 1G-4734
 20 SEQ ID NO: 82 is the determined cDNA sequence for 1H-4774
 SEQ ID NO: 83 is the determined cDNA sequence for 1H-4781
 SEQ ID NO: 84 is the determined cDNA sequence for 1H-4785
 SEQ ID NO: 85 is the determined cDNA sequence for 1H-4787
 SEQ ID NO: 86 is the determined cDNA sequence for 1H-4796
 25 SEQ ID NO: 87 is the determined cDNA sequence for 1I-4807
 SEQ ID NO: 88 is the determined cDNA sequence for 1I-4810
 SEQ ID NO: 89 is the determined cDNA sequence for 1I-4811
 SEQ ID NO: 90 is the determined cDNA sequence for 1J-4876
 SEQ ID NO: 91 is the determined cDNA sequence for 1K-4884

- SEQ ID NO: 92 is the determined cDNA sequence for 1K-4896
- SEQ ID NO: 93 is the determined cDNA sequence for 1G-4761
- SEQ ID NO: 94 is the determined cDNA sequence for 1G-4762
- SEQ ID NO: 95 is the determined cDNA sequence for 1H-4766
- 5 SEQ ID NO: 96 is the determined cDNA sequence for 1H-4770
- SEQ ID NO: 97 is the determined cDNA sequence for 1H-4771
- SEQ ID NO: 98 is the determined cDNA sequence for 1H-4772
- SEQ ID NO: 99 is the determined cDNA sequence for 1D-4297
- SEQ ID NO: 100 is the determined cDNA sequence for 1D-4309
- 10 SEQ ID NO: 101 is the determined cDNA sequence for 1D.1-4278
- SEQ ID NO: 102 is the determined cDNA sequence for 1D-4288
- SEQ ID NO: 103 is the determined cDNA sequence for 1D-4283
- SEQ ID NO: 104 is the determined cDNA sequence for 1D-4304
- SEQ ID NO: 105 is the determined cDNA sequence for 1D-4296
- 15 SEQ ID NO: 106 is the determined cDNA sequence for 1D-4280
- SEQ ID NO: 107 is the determined full length cDNA sequence for F1-12
(also referred to as P504S)
- SEQ ID NO: 108 is the predicted amino acid sequence for F1-12
- SEQ ID NO: 109 is the determined full length cDNA sequence for J1-17
- 20 SEQ ID NO: 110 is the determined full length cDNA sequence for L1-12
(also referred to as P501S)
- SEQ ID NO: 111 is the determined full length cDNA sequence for N1-1862
(also referred to as P503S)
- SEQ ID NO: 112 is the predicted amino acid sequence for J1-17
- 25 SEQ ID NO: 113 is the predicted amino acid sequence for L1-12 (also
referred to as P501S)
- SEQ ID NO: 114 is the predicted amino acid sequence for N1-1862 (also
referred to as P503S)
- SEQ ID NO: 115 is the determined cDNA sequence for P89

SEQ ID NO: 116 is the determined cDNA sequence for P90
 SEQ ID NO: 117 is the determined cDNA sequence for P92
 SEQ ID NO: 118 is the determined cDNA sequence for P95
 SEQ ID NO: 119 is the determined cDNA sequence for P98
 5 SEQ ID NO: 120 is the determined cDNA sequence for P102
 SEQ ID NO: 121 is the determined cDNA sequence for P110
 SEQ ID NO: 122 is the determined cDNA sequence for P111
 SEQ ID NO: 123 is the determined cDNA sequence for P114
 SEQ ID NO: 124 is the determined cDNA sequence for P115
 10 SEQ ID NO: 125 is the determined cDNA sequence for P116
 SEQ ID NO: 126 is the determined cDNA sequence for P124
 SEQ ID NO: 127 is the determined cDNA sequence for P126
 SEQ ID NO: 128 is the determined cDNA sequence for P130
 SEQ ID NO: 129 is the determined cDNA sequence for P133
 15 SEQ ID NO: 130 is the determined cDNA sequence for P138
 SEQ ID NO: 131 is the determined cDNA sequence for P143
 SEQ ID NO: 132 is the determined cDNA sequence for P151
 SEQ ID NO: 133 is the determined cDNA sequence for P156
 SEQ ID NO: 134 is the determined cDNA sequence for P157
 20 SEQ ID NO: 135 is the determined cDNA sequence for P166
 SEQ ID NO: 136 is the determined cDNA sequence for P176
 SEQ ID NO: 137 is the determined cDNA sequence for P178
 SEQ ID NO: 138 is the determined cDNA sequence for P179
 SEQ ID NO: 139 is the determined cDNA sequence for P185
 25 SEQ ID NO: 140 is the determined cDNA sequence for P192
 SEQ ID NO: 141 is the determined cDNA sequence for P201
 SEQ ID NO: 142 is the determined cDNA sequence for P204
 SEQ ID NO: 143 is the determined cDNA sequence for P208
 SEQ ID NO: 144 is the determined cDNA sequence for P211

SEQ ID NO: 145 is the determined cDNA sequence for P213
 SEQ ID NO: 146 is the determined cDNA sequence for P219
 SEQ ID NO: 147 is the determined cDNA sequence for P237
 SEQ ID NO: 148 is the determined cDNA sequence for P239
 5 SEQ ID NO: 149 is the determined cDNA sequence for P248
 SEQ ID NO: 150 is the determined cDNA sequence for P251
 SEQ ID NO: 151 is the determined cDNA sequence for P255
 SEQ ID NO: 152 is the determined cDNA sequence for P256
 SEQ ID NO: 153 is the determined cDNA sequence for P259
 10 SEQ ID NO: 154 is the determined cDNA sequence for P260
 SEQ ID NO: 155 is the determined cDNA sequence for P263
 SEQ ID NO: 156 is the determined cDNA sequence for P264
 SEQ ID NO: 157 is the determined cDNA sequence for P266
 SEQ ID NO: 158 is the determined cDNA sequence for P270
 15 SEQ ID NO: 159 is the determined cDNA sequence for P272
 SEQ ID NO: 160 is the determined cDNA sequence for P278
 SEQ ID NO: 161 is the determined cDNA sequence for P105
 SEQ ID NO: 162 is the determined cDNA sequence for P107
 SEQ ID NO: 163 is the determined cDNA sequence for P137
 20 SEQ ID NO: 164 is the determined cDNA sequence for P194
 SEQ ID NO: 165 is the determined cDNA sequence for P195
 SEQ ID NO: 166 is the determined cDNA sequence for P196
 SEQ ID NO: 167 is the determined cDNA sequence for P220
 SEQ ID NO: 168 is the determined cDNA sequence for P234
 25 SEQ ID NO: 169 is the determined cDNA sequence for P235
 SEQ ID NO: 170 is the determined cDNA sequence for P243
 SEQ ID NO: 171 is the determined cDNA sequence for P703P-DE1
 SEQ ID NO: 172 is the predicted amino acid sequence for P703P-DE1
 SEQ ID NO: 173 is the determined cDNA sequence for P703P-DE2

SEQ ID NO: 174 is the determined cDNA sequence for P703P-DE6
 SEQ ID NO: 175 is the determined cDNA sequence for P703P-DE13
 SEQ ID NO: 176 is the predicted amino acid sequence for P703P-DE13
 SEQ ID NO: 177 is the determined cDNA sequence for P703P-DE14
 5 SEQ ID NO: 178 is the predicted amino acid sequence for P703P-DE14
 SEQ ID NO: 179 is the determined extended cDNA sequence for 1G-4736
 SEQ ID NO: 180 is the determined extended cDNA sequence for 1G-4738
 SEQ ID NO: 181 is the determined extended cDNA sequence for 1G-4741
 SEQ ID NO: 182 is the determined extended cDNA sequence for 1G-4744
 10 SEQ ID NO: 183 is the determined extended cDNA sequence for 1H-4774
 SEQ ID NO: 184 is the determined extended cDNA sequence for 1H-4781
 SEQ ID NO: 185 is the determined extended cDNA sequence for 1H-4785
 SEQ ID NO: 186 is the determined extended cDNA sequence for 1H-4787
 SEQ ID NO: 187 is the determined extended cDNA sequence for 1H-4796
 15 SEQ ID NO: 188 is the determined extended cDNA sequence for 1I-4807
 SEQ ID NO: 189 is the determined 3' cDNA sequence for 1I-4810
 SEQ ID NO: 190 is the determined 3' cDNA sequence for 1I-4811
 SEQ ID NO: 191 is the determined extended cDNA sequence for 1J-4876
 SEQ ID NO: 192 is the determined extended cDNA sequence for 1K-4884
 20 SEQ ID NO: 193 is the determined extended cDNA sequence for 1K-4896
 SEQ ID NO: 194 is the determined extended cDNA sequence for 1G-4761
 SEQ ID NO: 195 is the determined extended cDNA sequence for 1G-4762
 SEQ ID NO: 196 is the determined extended cDNA sequence for 1H-4766
 SEQ ID NO: 197 is the determined 3' cDNA sequence for 1H-4770
 25 SEQ ID NO: 198 is the determined 3' cDNA sequence for 1H-4771
 SEQ ID NO: 199 is the determined extended cDNA sequence for 1H-4772
 SEQ ID NO: 200 is the determined extended cDNA sequence for 1D-4309
 SEQ ID NO: 201 is the determined extended cDNA sequence for 1D.1-4278
 SEQ ID NO: 202 is the determined extended cDNA sequence for 1D-4288

SEQ ID NO: 203 is the determined extended cDNA sequence for 1D-4283
 SEQ ID NO: 204 is the determined extended cDNA sequence for 1D-4304
 SEQ ID NO: 205 is the determined extended cDNA sequence for 1D-4296
 SEQ ID NO: 206 is the determined extended cDNA sequence for 1D-4280
 5 SEQ ID NO: 207 is the determined cDNA sequence for 10-d8fwd
 SEQ ID NO: 208 is the determined cDNA sequence for 10-H10con
 SEQ ID NO: 209 is the determined cDNA sequence for 11-C8rev
 SEQ ID NO: 210 is the determined cDNA sequence for 7.g6fwd
 SEQ ID NO: 211 is the determined cDNA sequence for 7.g6rev
 10 SEQ ID NO: 212 is the determined cDNA sequence for 8-b5fwd
 SEQ ID NO: 213 is the determined cDNA sequence for 8-b5rev
 SEQ ID NO: 214 is the determined cDNA sequence for 8-b6fwd
 SEQ ID NO: 215 is the determined cDNA sequence for 8-b6 rev
 SEQ ID NO: 216 is the determined cDNA sequence for 8-d4fwd
 15 SEQ ID NO: 217 is the determined cDNA sequence for 8-d9rev
 SEQ ID NO: 218 is the determined cDNA sequence for 8-g3fwd
 SEQ ID NO: 219 is the determined cDNA sequence for 8-g3rev
 SEQ ID NO: 220 is the determined cDNA sequence for 8-h11rev
 SEQ ID NO: 221 is the determined cDNA sequence for g-f12fwd
 20 SEQ ID NO: 222 is the determined cDNA sequence for g-f3rev
 SEQ ID NO: 223 is the determined cDNA sequence for P509S
 SEQ ID NO: 224 is the determined cDNA sequence for P510S
 SEQ ID NO: 225 is the determined cDNA sequence for P703DE5
 SEQ ID NO: 226 is the determined cDNA sequence for 9-A11
 25 SEQ ID NO: 227 is the determined cDNA sequence for 8-C6
 SEQ ID NO: 228 is the determined cDNA sequence for 8-H7
 SEQ ID NO: 229 is the determined cDNA sequence for JPTPN13
 SEQ ID NO: 230 is the determined cDNA sequence for JPTPN14
 SEQ ID NO: 231 is the determined cDNA sequence for JPTPN23

SEQ ID NO: 232 is the determined cDNA sequence for JPTPN24
 SEQ ID NO: 233 is the determined cDNA sequence for JPTPN25
 SEQ ID NO: 234 is the determined cDNA sequence for JPTPN30
 SEQ ID NO: 235 is the determined cDNA sequence for JPTPN34
 5 SEQ ID NO: 236 is the determined cDNA sequence for PTPN35
 SEQ ID NO: 237 is the determined cDNA sequence for JPTPN36
 SEQ ID NO: 238 is the determined cDNA sequence for JPTPN38
 SEQ ID NO: 239 is the determined cDNA sequence for JPTPN39
 SEQ ID NO: 240 is the determined cDNA sequence for JPTPN40
 10 SEQ ID NO: 241 is the determined cDNA sequence for JPTPN41
 SEQ ID NO: 242 is the determined cDNA sequence for JPTPN42
 SEQ ID NO: 243 is the determined cDNA sequence for JPTPN45
 SEQ ID NO: 244 is the determined cDNA sequence for JPTPN46
 SEQ ID NO: 245 is the determined cDNA sequence for JPTPN51
 15 SEQ ID NO: 246 is the determined cDNA sequence for JPTPN56
 SEQ ID NO: 247 is the determined cDNA sequence for PTPN64
 SEQ ID NO: 248 is the determined cDNA sequence for JPTPN65
 SEQ ID NO: 249 is the determined cDNA sequence for JPTPN67
 SEQ ID NO: 250 is the determined cDNA sequence for JPTPN76
 20 SEQ ID NO: 251 is the determined cDNA sequence for JPTPN84
 SEQ ID NO: 252 is the determined cDNA sequence for JPTPN85
 SEQ ID NO: 253 is the determined cDNA sequence for JPTPN86
 SEQ ID NO: 254 is the determined cDNA sequence for JPTPN87
 SEQ ID NO: 255 is the determined cDNA sequence for JPTPN88
 25 SEQ ID NO: 256 is the determined cDNA sequence for JP1F1
 SEQ ID NO: 257 is the determined cDNA sequence for JP1F2
 SEQ ID NO: 258 is the determined cDNA sequence for JP1C2
 SEQ ID NO: 259 is the determined cDNA sequence for JP1B1
 SEQ ID NO: 260 is the determined cDNA sequence for JP1B2

SEQ ID NO: 261 is the determined cDNA sequence for JP1D3
 SEQ ID NO: 262 is the determined cDNA sequence for JP1A4
 SEQ ID NO: 263 is the determined cDNA sequence for JP1F5
 SEQ ID NO: 264 is the determined cDNA sequence for JP1E6
 5 SEQ ID NO: 265 is the determined cDNA sequence for JP1D6
 SEQ ID NO: 266 is the determined cDNA sequence for JP1B5
 SEQ ID NO: 267 is the determined cDNA sequence for JP1A6
 SEQ ID NO: 268 is the determined cDNA sequence for JP1E8
 SEQ ID NO: 269 is the determined cDNA sequence for JP1D7
 10 SEQ ID NO: 270 is the determined cDNA sequence for JP1D9
 SEQ ID NO: 271 is the determined cDNA sequence for JP1C10
 SEQ ID NO: 272 is the determined cDNA sequence for JP1A9
 SEQ ID NO: 273 is the determined cDNA sequence for JP1F12
 SEQ ID NO: 274 is the determined cDNA sequence for JP1E12
 15 SEQ ID NO: 275 is the determined cDNA sequence for JP1D11
 SEQ ID NO: 276 is the determined cDNA sequence for JP1C11
 SEQ ID NO: 277 is the determined cDNA sequence for JP1C12
 SEQ ID NO: 278 is the determined cDNA sequence for JP1B12
 SEQ ID NO: 279 is the determined cDNA sequence for JP1A12
 20 SEQ ID NO: 280 is the determined cDNA sequence for JP8G2
 SEQ ID NO: 281 is the determined cDNA sequence for JP8H1
 SEQ ID NO: 282 is the determined cDNA sequence for JP8H2
 SEQ ID NO: 283 is the determined cDNA sequence for JP8A3
 SEQ ID NO: 284 is the determined cDNA sequence for JP8A4
 25 SEQ ID NO: 285 is the determined cDNA sequence for JP8C3
 SEQ ID NO: 286 is the determined cDNA sequence for JP8G4
 SEQ ID NO: 287 is the determined cDNA sequence for JP8B6
 SEQ ID NO: 288 is the determined cDNA sequence for JP8D6
 SEQ ID NO: 289 is the determined cDNA sequence for JP8F5

SEQ ID NO: 290 is the determined cDNA sequence for JP8A8
 SEQ ID NO: 291 is the determined cDNA sequence for JP8C7
 SEQ ID NO: 292 is the determined cDNA sequence for JP8D7
 SEQ ID NO: 293 is the determined cDNA sequence for P8D8
 5 SEQ ID NO: 294 is the determined cDNA sequence for JP8E7
 SEQ ID NO: 295 is the determined cDNA sequence for JP8F8
 SEQ ID NO: 296 is the determined cDNA sequence for JP8G8
 SEQ ID NO: 297 is the determined cDNA sequence for JP8B10
 SEQ ID NO: 298 is the determined cDNA sequence for JP8C10
 10 SEQ ID NO: 299 is the determined cDNA sequence for JP8E9
 SEQ ID NO: 300 is the determined cDNA sequence for JP8E10
 SEQ ID NO: 301 is the determined cDNA sequence for JP8F9
 SEQ ID NO: 302 is the determined cDNA sequence for JP8H9
 SEQ ID NO: 303 is the determined cDNA sequence for JP8C12
 15 SEQ ID NO: 304 is the determined cDNA sequence for JP8E11
 SEQ ID NO: 305 is the determined cDNA sequence for JP8E12
 SEQ ID NO: 306 is the amino acid sequence for the peptide PS2#12
 SEQ ID NO: 307 is the determined cDNA sequence for P711P
 SEQ ID NO: 308 is the determined cDNA sequence for P712P
 20 SEQ ID NO: 309 is the determined cDNA sequence for CLONE23
 SEQ ID NO: 310 is the determined cDNA sequence for P774P
 SEQ ID NO: 311 is the determined cDNA sequence for P775P
 SEQ ID NO: 312 is the determined cDNA sequence for P715P
 SEQ ID NO: 313 is the determined cDNA sequence for P710P
 25 SEQ ID NO: 314 is the determined cDNA sequence for P767P
 SEQ ID NO: 315 is the determined cDNA sequence for P768P
 SEQ ID NO: 316-325 are the determined cDNA sequences of previously
 isolated genes
 SEQ ID NO: 326 is the determined cDNA sequence for P703PDE5

- SEQ ID NO: 327 is the predicted amino acid sequence for P703PDE5
 SEQ ID NO: 328 is the determined cDNA sequence for P703P6.26
 SEQ ID NO: 329 is the predicted amino acid sequence for P703P6.26
 SEQ ID NO: 330 is the determined cDNA sequence for P703PX-23
 5 SEQ ID NO: 331 is the predicted amino acid sequence for P703PX-23
 SEQ ID NO: 332 is the determined full length cDNA sequence for P509S
 SEQ ID NO: 333 is the determined extended cDNA sequence for P707P
 (also referred to as 11-C9)
 SEQ ID NO: 334 is the determined cDNA sequence for P714P
 10 SEQ ID NO: 335 is the determined cDNA sequence for P705P (also
 referred to as 9-F3)
 SEQ ID NO: 336 is the predicted amino acid sequence for P705P
 SEQ ID NO: 337 is the amino acid sequence of the peptide P1S#10
 SEQ ID NO: 338 is the amino acid sequence of the peptide p5
 15 SEQ ID NO: 339 is the predicted amino acid sequence of P509S
 SEQ ID NO: 340 is the determined cDNA sequence for P778P
 SEQ ID NO: 341 is the determined cDNA sequence for P786P
 SEQ ID NO: 342 is the determined cDNA sequence for P789P
 SEQ ID NO: 343 is the determined cDNA sequence for a clone showing
 20 homology to Homo sapiens MM46 mRNA
 SEQ ID NO: 344 is the determined cDNA sequence for a clone showing
 homology to Homo sapiens TNF-alpha stimulated ABC protein (ABC50) mRNA
 SEQ ID NO: 345 is the determined cDNA sequence for a clone showing
 homology to Homo sapiens mRNA for E-cadherin
 25 SEQ ID NO: 346 is the determined cDNA sequence for a clone showing
 homology to Human nuclear-encoded mitochondrial serine hydroxymethyltransferase
 (SHMT)
 SEQ ID NO: 347 is the determined cDNA sequence for a clone showing
 homology to Homo sapiens natural resistance-associated macrophage protein2 (NRAMP2)

SEQ ID NO: 348 is the determined cDNA sequence for a clone showing homology to Homo sapiens phosphoglucomutase-related protein (PGMRP)

SEQ ID NO: 349 is the determined cDNA sequence for a clone showing homology to Human mRNA for proteosome subunit p40

- 5 SEQ ID NO: 350 is the determined cDNA sequence for P777P
 SEQ ID NO: 351 is the determined cDNA sequence for P779P
 SEQ ID NO: 352 is the determined cDNA sequence for P790P
 SEQ ID NO: 353 is the determined cDNA sequence for P784P
 SEQ ID NO: 354 is the determined cDNA sequence for P776P
 10 SEQ ID NO: 355 is the determined cDNA sequence for P780P
 SEQ ID NO: 356 is the determined cDNA sequence for P544S
 SEQ ID NO: 357 is the determined cDNA sequence for P745S
 SEQ ID NO: 358 is the determined cDNA sequence for P782P
 SEQ ID NO: 359 is the determined cDNA sequence for P783P
 15 SEQ ID NO: 360 is the determined cDNA sequence for unknown 17984
 SEQ ID NO: 361 is the determined cDNA sequence for P787P
 SEQ ID NO: 362 is the determined cDNA sequence for P788P
 SEQ ID NO: 363 is the determined cDNA sequence for unknown 17994
 SEQ ID NO: 364 is the determined cDNA sequence for P781P
 20 SEQ ID NO: 365 is the determined cDNA sequence for P785P
 SEQ ID NO: 366-375 are the determined cDNA sequences for splice variants of B305D.

SEQ ID NO: 376 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 366.

- 25 SEQ ID NO: 377 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 372.

SEQ ID NO: 378 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 373.

SEQ ID NO: 379 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 374.

SEQ ID NO: 380 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 375.

- 5 SEQ ID NO: 381 is the determined cDNA sequence for B716P.
 SEQ ID NO: 382 is the determined full-length cDNA sequence for P711P.
 SEQ ID NO: 383 is the predicted amino acid sequence for P711P.
 SEQ ID NO: 384 is the cDNA sequence for P1000C.
 SEQ ID NO: 385 is the cDNA sequence for CGI-82.
- 10 SEQ ID NO:386 is the cDNA sequence for 23320.
 SEQ ID NO:387 is the cDNA sequence for CGI-69.
 SEQ ID NO:388 is the cDNA sequence for L-iditol-2-dehydrogenase.
 SEQ ID NO:389 is the cDNA sequence for 23379.
 SEQ ID NO:390 is the cDNA sequence for 23381.
- 15 SEQ ID NO:391 is the cDNA sequence for KIAA0122.
 SEQ ID NO:392 is the cDNA sequence for 23399.
 SEQ ID NO:393 is the cDNA sequence for a previously identified gene.
 SEQ ID NO:394 is the cDNA sequence for HCLBP.
 SEQ ID NO:395 is the cDNA sequence for transglutaminase.
- 20 SEQ ID NO:396 is the cDNA sequence for a previously identified gene.
 SEQ ID NO:397 is the cDNA sequence for PAP.
 SEQ ID NO:398 is the cDNA sequence for Ets transcription factor PDEF.
 SEQ ID NO:399 is the cDNA sequence for hTGR.
 SEQ ID NO:400 is the cDNA sequence for KIAA0295.
- 25 SEQ ID NO:401 is the cDNA sequence for 22545.
 SEQ ID NO:402 is the cDNA sequence for 22547.
 SEQ ID NO:403 is the cDNA sequence for 22548.
 SEQ ID NO:404 is the cDNA sequence for 22550.
 SEQ ID NO:405 is the cDNA sequence for 22551.

SEQ ID NO:406 is the cDNA sequence for 22552.

SEQ ID NO:407 is the cDNA sequence for 22553 (also known as P1020C).

SEQ ID NO:408 is the cDNA sequence for 22558.

SEQ ID NO:409 is the cDNA sequence for 22562.

5 SEQ ID NO:410 is the cDNA sequence for 22565.

SEQ ID NO:411 is the cDNA sequence for 22567.

SEQ ID NO:412 is the cDNA sequence for 22568.

SEQ ID NO:413 is the cDNA sequence for 22570.

SEQ ID NO:414 is the cDNA sequence for 22571.

10 SEQ ID NO:415 is the cDNA sequence for 22572.

SEQ ID NO:416 is the cDNA sequence for 22573.

SEQ ID NO:417 is the cDNA sequence for 22573.

SEQ ID NO:418 is the cDNA sequence for 22575.

SEQ ID NO:419 is the cDNA sequence for 22580.

15 SEQ ID NO:420 is the cDNA sequence for 22581.

SEQ ID NO:421 is the cDNA sequence for 22582.

SEQ ID NO:422 is the cDNA sequence for 22583.

SEQ ID NO:423 is the cDNA sequence for 22584.

SEQ ID NO:424 is the cDNA sequence for 22585.

20 SEQ ID NO:425 is the cDNA sequence for 22586.

SEQ ID NO:426 is the cDNA sequence for 22587.

SEQ ID NO:427 is the cDNA sequence for 22588.

SEQ ID NO:428 is the cDNA sequence for 22589.

SEQ ID NO:429 is the cDNA sequence for 22590.

25 SEQ ID NO:430 is the cDNA sequence for 22591.

SEQ ID NO:431 is the cDNA sequence for 22592.

SEQ ID NO:432 is the cDNA sequence for 22593.

SEQ ID NO:433 is the cDNA sequence for 22594.

SEQ ID NO:434 is the cDNA sequence for 22595.

SEQ ID NO:435 is the cDNA sequence for 22596.
 SEQ ID NO:436 is the cDNA sequence for 22847.
 SEQ ID NO:437 is the cDNA sequence for 22848.
 SEQ ID NO:438 is the cDNA sequence for 22849.
 5 SEQ ID NO:439 is the cDNA sequence for 22851.
 SEQ ID NO:440 is the cDNA sequence for 22852.
 SEQ ID NO:441 is the cDNA sequence for 22853.
 SEQ ID NO:442 is the cDNA sequence for 22854.
 SEQ ID NO:443 is the cDNA sequence for 22855.
 10 SEQ ID NO:444 is the cDNA sequence for 22856.
 SEQ ID NO:445 is the cDNA sequence for 22857.
 SEQ ID NO:446 is the cDNA sequence for 23601.
 SEQ ID NO:447 is the cDNA sequence for 23602.
 SEQ ID NO:448 is the cDNA sequence for 23605.
 15 SEQ ID NO:449 is the cDNA sequence for 23606.
 SEQ ID NO:450 is the cDNA sequence for 23612.
 SEQ ID NO:451 is the cDNA sequence for 23614.
 SEQ ID NO:452 is the cDNA sequence for 23618.
 SEQ ID NO:453 is the cDNA sequence for 23622.
 20 SEQ ID NO:454 is the cDNA sequence for folate hydrolase.
 SEQ ID NO:455 is the cDNA sequence for LIM protein.
 SEQ ID NO:456 is the cDNA sequence for a known gene.
 SEQ ID NO:457 is the cDNA sequence for a known gene.
 SEQ ID NO:458 is the cDNA sequence for a previously identified gene.
 25 SEQ ID NO:459 is the cDNA sequence for 23045.
 SEQ ID NO:460 is the cDNA sequence for 23032.
 SEQ ID NO:461 is the cDNA sequence for 23054.
 SEQ ID NO:462-467 are cDNA sequences for known genes.
 SEQ ID NO:468-471 are cDNA sequences for P710P.

SEQ ID NO:472 is a cDNA sequence for P1001C.

SEQ ID NO: 473 is the determined cDNA sequence for a first splice variant of P775P (referred to as 27505).

5 SEQ ID NO: 474 is the determined cDNA sequence for a second splice variant of P775P (referred to as 19947).

SEQ ID NO: 475 is the determined cDNA sequence for a third splice variant of P775P (referred to as 19941).

SEQ ID NO: 476 is the determined cDNA sequence for a fourth splice variant of P775P (referred to as 19937).

10 SEQ ID NO: 477 is a first predicted amino acid sequence encoded by the sequence of SEQ ID NO: 474.

SEQ ID NO: 478 is a second predicted amino acid sequence encoded by the sequence of SEQ ID NO: 474.

15 SEQ ID NO: 479 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: 475.

SEQ ID NO: 480 is a first predicted amino acid sequence encoded by the sequence of SEQ ID NO: 473.

SEQ ID NO: 481 is a second predicted amino acid sequence encoded by the sequence of SEQ ID NO: 473.

20 SEQ ID NO: 482 is a third predicted amino acid sequence encoded by the sequence of SEQ ID NO: 473.

SEQ ID NO: 483 is a fourth predicted amino acid sequence encoded by the sequence of SEQ ID NO: 473.

25 SEQ ID NO: 484 is the first 30 amino acids of the *M. tuberculosis* antigen Ra12.

SEQ ID NO: 485 is the PCR primer AW025.

SEQ ID NO: 486 is the PCR primer AW003.

SEQ ID NO: 487 is the PCR primer AW027.

SEQ ID NO: 488 is the PCR primer AW026.

SEQ ID NO: 489-501 are peptides employed in epitope mapping studies.

SEQ ID NO: 502 is the determined cDNA sequence of the complementarity determining region for the anti-P503S monoclonal antibody 20D4.

5 SEQ ID NO: 503 is the determined cDNA sequence of the complementarity determining region for the anti-P503S monoclonal antibody JA1.

SEQ ID NO: 504 & 505 are peptides employed in epitope mapping studies.

SEQ ID NO: 506 is the determined cDNA sequence of the complementarity determining region for the anti-P703P monoclonal antibody 8H2.

10 SEQ ID NO: 507 is the determined cDNA sequence of the complementarity determining region for the anti-P703P monoclonal antibody 7H8.

SEQ ID NO: 508 is the determined cDNA sequence of the complementarity determining region for the anti-P703P monoclonal antibody 2D4.

SEQ ID NO: 509-522 are peptides employed in epitope mapping studies.

15 SEQ ID NO: 523 is a mature form of P703P used to raise antibodies against P703P.

SEQ ID NO: 524 is the putative full-length cDNA sequence of P703P.

SEQ ID NO: 525 is the predicted amino acid sequence encoded by SEQ ID NO: 524.

SEQ ID NO: 526 is the full-length cDNA sequence for P790P.

20 SEQ ID NO: 527 is the predicted amino acid sequence for P790P.

SEQ ID NO: 528 & 529 are PCR primers.

SEQ ID NO: 530 is the cDNA sequence of a splice variant of SEQ ID NO: 366.

25 SEQ ID NO: 531 is the cDNA sequence of the open reading frame of SEQ ID NO: 530.

SEQ ID NO: 532 is the predicted amino acid encoded by the sequence of SEQ ID NO: 531.

SEQ ID NO: 533 is the DNA sequence of a putative ORF of P775P.

SEQ ID NO: 534 is the predicted amino acid sequence encoded by SEQ ID NO: 533.

SEQ ID NO: 535 is a first full-length cDNA sequence for P510S.

SEQ ID NO: 536 is a second full-length cDNA sequence for P510S.

5 SEQ ID NO: 537 is the predicted amino acid sequence encoded by SEQ ID NO: 535.

SEQ ID NO: 538 is the predicted amino acid sequence encoded by SEQ ID NO: 536.

SEQ ID NO: 539 is the peptide P501S-370.

10 SEQ ID NO: 540 is the peptide P501S-376.

SEQ ID NO: 541-551 are epitopes of P501S.

SEQ ID NO: 552 is an extended cDNA sequence for P712P.

SEQ ID NO: 553-568 are the amino acid sequences encoded by predicted open reading frames within SEQ ID NO: 552.

15 SEQ ID NO: 569 is an extended cDNA sequence for P776P.

SEQ ID NO: 570 is the determined cDNA sequence for a splice variant of P776P referred to as contig 6.

SEQ ID NO: 571 is the determined cDNA sequence for a splice variant of P776P referred to as contig 7.

20 SEQ ID NO: 572 is the determined cDNA sequence for a splice variant of P776P referred to as contig 14.

SEQ ID NO: 573 is the amino acid sequence encoded by a first predicted ORF of SEQ ID NO: 570.

25 SEQ ID NO: 574 is the amino acid sequence encoded by a second predicted ORF of SEQ ID NO: 570.

SEQ ID NO: 575 is the amino acid sequence encoded by a predicted ORF of SEQ ID NO: 571.

SEQ ID NO: 576-586 are amino acid sequences encoded by predicted ORFs of SEQ ID NO: 569.

SEQ ID NO: 587 is a DNA consensus sequence of the sequences of P767P and P777P.

SEQ ID NO: 588-590 are amino acid sequences encoded by predicted ORFs of SEQ ID NO: 587.

5 SEQ ID NO: 591 is an extended cDNA sequence for P1020C.

SEQ ID NO: 592 is the predicted amino acid sequence encoded by the sequence of SEQ ID NO: P1020C.

SEQ ID NO: 593 is a splice variant of P775P referred to as 50748.

10 SEQ ID NO: 594 is a splice variant of P775P referred to as 50717. SEQ ID NO: 595 is a splice variant of P775P referred to as 45985.

SEQ ID NO: 596 is a splice variant of P775P referred to as 38769.

SEQ ID NO: 597 is a splice variant of P775P referred to as 37922.

SEQ ID NO: 598 is a splice variant of P510S referred to as 49274.

SEQ ID NO: 599 is a splice variant of P510S referred to as 39487.

15 SEQ ID NO: 600 is a splice variant of P504S referred to as 5167.16.

SEQ ID NO: 601 is a splice variant of P504S referred to as 5167.1.

SEQ ID NO: 602 is a splice variant of P504S referred to as 5163.46.

SEQ ID NO: 603 is a splice variant of P504S referred to as 5163.42.

SEQ ID NO: 604 is a splice variant of P504S referred to as 5163.34.

20 SEQ ID NO: 605 is a splice variant of P504S referred to as 5163.17.

SEQ ID NO: 606 is a splice variant of P501S referred to as 10640.

SEQ ID NO: 607-615 are the sequences of PCR primers.

SEQ ID NO: 616 is the determined cDNA sequence of a fusion of P703P and PSA.

25 SEQ ID NO: 617 is the amino acid sequence of the fusion of P703P and PSA.

SEQ ID NO: 618-689 are determined cDNA sequences of prostate-specific clones.

SEQ ID NO: 690 is the cDNA sequence of the gene DD3.

SEQ ID NO: 691-697 are determined cDNA sequences of prostate-specific clones.

SEQ ID NO: 698 is an extended cDNA sequence for P714P.

SEQ ID NO: 699-701 are the cDNA sequences for splice variants of P704P.

5 SEQ ID NO: 702 is the cDNA sequence of a spliced variant of P553S referred to as P553S-14.

SEQ ID NO: 703 is the cDNA sequence of a spliced variant of P553S referred to as P553S-12.

10 SEQ ID NO: 704 is the cDNA sequence of a spliced variant of P553S referred to as P553S-10.

SEQ ID NO: 705 is the cDNA sequence of a spliced variant of P553S referred to as P553S-6.

SEQ ID NO: 706 is the amino acid sequence encoded by SEQ ID NO: 705.

SEQ ID NO: 707 is the amino acid sequence encoded by SEQ ID NO: 702

15 SEQ ID NO: 708 is a second amino acid sequence encoded by SEQ ID NO: 702.

SEQ ID NO: 709-772 are determined cDNA sequences of prostate-specific clones.

SEQ ID NO: 773 is a first full-length cDNA sequence for prostate-specific transglutaminase gene (also referred to herein as P558S).

20 SEQ ID NO: 774 is a second full-length cDNA sequence for prostate-specific transglutaminase gene.

SEQ ID NO: 775 is the amino acid sequence encoded by the sequence of SEQ ID NO: 773.

25 SEQ ID NO: 776 is the amino acid sequence encoded by the sequence of SEQ ID NO: 774.

SEQ ID NO: 777 is the full-length cDNA sequence for P788P.

SEQ ID NO: 778 is the amino acid sequence encoded by SEQ ID NO: 777.

SEQ ID NO: 779 is the determined cDNA sequence for a polymorphic variant of P788P.

- SEQ ID NO: 780 is the amino acid sequence encoded by SEQ ID NO: 779.
 SEQ ID NO: 781 is the amino acid sequence of peptide 4 from P703P.
 SEQ ID NO: 782 is the cDNA sequence that encodes peptide 4 from P703P.
 SEQ ID NO: 783-798 are the cDNA sequence encoding epitopes of P703P.
 5 SEQ ID NO: 799-814 are the amino acid sequences of epitopes of P703P.
 SEQ ID NO: 815 and 816 are PCR primers.
 SEQ ID NO: 817 is the cDNA sequence encoding an N-terminal portion of
 P788P expressed in *E. coli*.
 SEQ ID NO: 818 is the amino acid sequence of the N-terminal portion of
 10 P788P expressed in *E. coli*.
 SEQ ID NO: 819 is the amino acid sequence of the *M. tuberculosis* antigen
 Ra12.
 SEQ ID NO: 820 and 821 are PCR primers.
 SEQ ID NO: 822 is the cDNA sequence for the Ra12-P510S-C construct.
 15 SEQ ID NO: 823 is the cDNA sequence for the P510S-C construct.
 SEQ ID NO: 824 is the cDNA sequence for the P510S-E3 construct.
 SEQ ID NO: 825 is the amino acid sequence for the Ra12-P510S-C
 construct.
 SEQ ID NO: 826 is the amino acid sequence for the P510S-C construct.
 20 SEQ ID NO: 827 is the amino acid sequence for the P510S-E3 construct.
 SEQ ID NO: 828-833 are PCR primers.
 SEQ ID NO: 834 is the cDNA sequence of the construct Ra12-P775P-
 ORF3.
 SEQ ID NO: 835 is the amino acid sequence of the construct Ra12-P775P-
 25 ORF3.

DETAILED DESCRIPTION OF THE INVENTION

As noted above, the present invention is generally directed to compositions and methods for using the compositions, for example in the therapy and diagnosis of

cancer, such as prostate cancer. Certain illustrative compositions described herein include prostate-specific polypeptides, polynucleotides encoding such polypeptides, binding agents such as antibodies, antigen presenting cells (APCs) and/or immune system cells (*e.g.*, T cells). A "prostate-specific protein," as the term is used herein, refers generally to a protein that is expressed in prostate cells at a level that is at least two fold, and preferably at least five fold, greater than the level of expression in other normal tissues, as determined using a representative assay provided herein. Certain prostate-specific proteins are tumor proteins that react detectably (within an immunoassay, such as an ELISA or Western blot) with antisera of a patient afflicted with prostate cancer.

Therefore, in accordance with the above, and as described further below, the present invention provides illustrative polynucleotide compositions having sequences set forth in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824, illustrative polypeptide compositions having amino acid sequences set forth in SEQ ID NO: 112-114, 172, 176, 178, 327, 329, 331, 336, 339, 376-380, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534, 537-551, 553-568, 573-586, 588-590, 592, 706-708, 775, 776, 778, 780, 781, 811, 814, 818, 826 and 827, antibody compositions capable of binding such polypeptides, and numerous additional embodiments employing such compositions, for example in the detection, diagnosis and/or therapy of human prostate cancer.

POLYNUCLEOTIDE COMPOSITIONS

As used herein, the terms "DNA segment" and "polynucleotide" refer to a DNA molecule that has been isolated free of total genomic DNA of a particular species. Therefore, a DNA segment encoding a polypeptide refers to a DNA segment that contains one or more coding sequences yet is substantially isolated away from, or purified free from, total genomic DNA of the species from which the DNA segment is obtained. Included within the terms "DNA segment" and "polynucleotide" are DNA segments and smaller

fragments of such segments, and also recombinant vectors, including, for example, plasmids, cosmids, phagemids, phage, viruses, and the like.

As will be understood by those skilled in the art, the DNA segments of this invention can include genomic sequences, extra-genomic and plasmid-encoded sequences
 5 and smaller engineered gene segments that express, or may be adapted to express, proteins, polypeptides, peptides and the like. Such segments may be naturally isolated, or modified synthetically by the hand of man.

"Isolated," as used herein, means that a polynucleotide is substantially away from other coding sequences, and that the DNA segment does not contain large portions of
 10 unrelated coding DNA, such as large chromosomal fragments or other functional genes or polypeptide coding regions. Of course, this refers to the DNA segment as originally isolated, and does not exclude genes or coding regions later added to the segment by the hand of man.

As will be recognized by the skilled artisan, polynucleotides may be single-
 15 stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. RNA molecules include HnRNA molecules, which contain introns and correspond to a DNA molecule in a one-to-one manner, and mRNA molecules, which do not contain introns. Additional coding or non-coding sequences may, but need not, be present within a polynucleotide of the present invention, and a polynucleotide may,
 20 but need not, be linked to other molecules and/or support materials.

Polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes a prostate-specific protein or a portion thereof) or may comprise a variant, or a biological or antigenic functional equivalent of such a sequence. Polynucleotide variants may contain one or more substitutions, additions, deletions and/or
 25 insertions, as further described below, preferably such that the immunogenicity of the encoded polypeptide is not diminished, relative to a native tumor protein. The effect on the immunogenicity of the encoded polypeptide may generally be assessed as described herein. The term "variants" also encompasses homologous genes of xenogenic origin.

When comparing polynucleotide or polypeptide sequences, two sequences are said to be “identical” if the sequence of nucleotides or amino acids in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A “comparison window” as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins – Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenies pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS* 5:151-153; Myers, E.W. and Muller W. (1988) *CABIOS* 4:11-17; Robinson, E.D. (1971) *Comb. Theor* 11:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol.* 4:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad., Sci. USA* 80:726-730.

Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Add. APL. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity methods of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP,

BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

One preferred example of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0 algorithms, which are described in Altschul *et al.* (1977) *Nucl. Acids Res.* 25:3389-3402 and Altschul *et al.* (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent sequence identity for the polynucleotides and polypeptides of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. In one illustrative example, cumulative scores can be calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). For amino acid sequences, a scoring matrix can be used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915) alignments, (B) of 50, expectation (E) of 10, M=5, N=-4 and a comparison of both strands.

Preferably, the “percentage of sequence identity” is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical

nucleic acid bases or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

5 Therefore, the present invention encompasses polynucleotide and polypeptide sequences having substantial identity to the sequences disclosed herein, for example those comprising at least 50% sequence identity, preferably at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% or higher, sequence identity compared to a polynucleotide or polypeptide sequence of this invention using the
10 methods described herein, (*e.g.*, BLAST analysis using standard parameters, as described below). One skilled in this art will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide sequences by taking into account codon degeneracy, amino acid similarity, reading frame positioning and the like.

15 In additional embodiments, the present invention provides isolated polynucleotides and polypeptides comprising various lengths of contiguous stretches of sequence identical to or complementary to one or more of the sequences disclosed herein. For example, polynucleotides are provided by this invention that comprise at least about 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500 or 1000 or more contiguous nucleotides
20 of one or more of the sequences disclosed herein as well as all intermediate lengths there between. It will be readily understood that "intermediate lengths", in this context, means any length between the quoted values, such as 16, 17, 18, 19, *etc.*; 21, 22, 23, *etc.*; 30, 31, 32, *etc.*; 50, 51, 52, 53, *etc.*; 100, 101, 102, 103, *etc.*; 150, 151, 152, 153, *etc.*; including all integers through 200-500; 500-1,000, and the like.

25 The polynucleotides of the present invention, or fragments thereof, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost

any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol. For example, illustrative DNA segments with total lengths of about 10,000, about 5000, about 3000, about 2,000, about 1,000, about 500, about 200, about 100, about 50 base pairs in length, and the like,
 5 (including all intermediate lengths) are contemplated to be useful in many implementations of this invention.

In other embodiments, the present invention is directed to polynucleotides that are capable of hybridizing under moderately stringent conditions to a polynucleotide sequence provided herein, or a fragment thereof, or a complementary sequence thereof.
 10 Hybridization techniques are well known in the art of molecular biology. For purposes of illustration, suitable moderately stringent conditions for testing the hybridization of a polynucleotide of this invention with other polynucleotides include prewashing in a solution of 5 X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50°C-65°C, 5 X SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X
 15 and 0.2X SSC containing 0.1% SDS.

Moreover, it will be appreciated by those of ordinary skill in the art that, as a result of the degeneracy of the genetic code, there are many nucleotide sequences that encode a polypeptide as described herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides that
 20 vary due to differences in codon usage are specifically contemplated by the present invention. Further, alleles of the genes comprising the polynucleotide sequences provided herein are within the scope of the present invention. Alleles are endogenous genes that are altered as a result of one or more mutations, such as deletions, additions and/or substitutions of nucleotides. The resulting mRNA and protein may, but need not, have an
 25 altered structure or function. Alleles may be identified using standard techniques (such as hybridization, amplification and/or database sequence comparison).

PROBES AND PRIMERS

In other embodiments of the present invention, the polynucleotide sequences provided herein can be advantageously used as probes or primers for nucleic acid hybridization. As such, it is contemplated that nucleic acid segments that comprise a
 5 sequence region of at least about 15 nucleotide long contiguous sequence that has the same sequence as, or is complementary to, a 15 nucleotide long contiguous sequence disclosed herein will find particular utility. Longer contiguous identical or complementary sequences, *e.g.*, those of about 20, 30, 40, 50, 100, 200, 500, 1000 (including all intermediate lengths) and even up to full length sequences will also be of use in certain
 10 embodiments.

The ability of such nucleic acid probes to specifically hybridize to a sequence of interest will enable them to be of use in detecting the presence of complementary sequences in a given sample. However, other uses are also envisioned, such as the use of the sequence information for the preparation of mutant species primers,
 15 or primers for use in preparing other genetic constructions.

Polynucleotide molecules having sequence regions consisting of contiguous nucleotide stretches of 10-14, 15-20, 30, 50, or even of 100-200 nucleotides or so (including intermediate lengths as well), identical or complementary to a polynucleotide sequence disclosed herein, are particularly contemplated as hybridization probes for use in,
 20 *e.g.*, Southern and Northern blotting. This would allow a gene product, or fragment thereof, to be analyzed, both in diverse cell types and also in various bacterial cells. The total size of fragment, as well as the size of the complementary stretch(es), will ultimately depend on the intended use or application of the particular nucleic acid segment. Smaller fragments will generally find use in hybridization embodiments, wherein the length of the
 25 contiguous complementary region may be varied, such as between about 15 and about 100 nucleotides, but larger contiguous complementarity stretches may be used, according to the length complementary sequences one wishes to detect.

The use of a hybridization probe of about 15-25 nucleotides in length allows the formation of a duplex molecule that is both stable and selective. Molecules having

contiguous complementary sequences over stretches greater than 15 bases in length are generally preferred, though, in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of specific hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having gene-complementary stretches of
 5 15 to 25 contiguous nucleotides, or even longer where desired.

Hybridization probes may be selected from any portion of any of the sequences disclosed herein. All that is required is to review the sequence set forth in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591,
 10 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824, or to any continuous portion of the sequence, from about 15-25 nucleotides in length up to and including the full length sequence, that one wishes to utilize as a probe or primer. The choice of probe and primer sequences may be governed by various factors. For example, one may wish to employ primers from towards the termini of the total sequence.

15 Small polynucleotide segments or fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer. Also, fragments may be obtained by application of nucleic acid reproduction technology, such as the PCR™ technology of U. S. Patent 4,683,202 (incorporated herein by reference), by introducing selected sequences into
 20 recombinant vectors for recombinant production, and by other recombinant DNA techniques generally known to those of skill in the art of molecular biology.

The nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of the entire gene or gene fragments of interest. Depending on the application envisioned, one will typically desire to
 25 employ varying conditions of hybridization to achieve varying degrees of selectivity of probe towards target sequence. For applications requiring high selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids, *e.g.*, one will select relatively low salt and/or high temperature conditions, such as provided by a salt concentration of from about 0.02 M to about 0.15 M salt at temperatures of from about

50°C to about 70°C. Such selective conditions tolerate little, if any, mismatch between the probe and the template or target strand, and would be particularly suitable for isolating related sequences.

Of course, for some applications, for example, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template, less stringent (reduced stringency) hybridization conditions will typically be needed in order to allow formation of the heteroduplex. In these circumstances, one may desire to employ salt conditions such as those of from about 0.15 M to about 0.9 M salt, at temperatures ranging from about 20°C to about 55°C. Cross-hybridizing species can thereby be readily identified as positively hybridizing signals with respect to control hybridizations. In any case, it is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide, which serves to destabilize the hybrid duplex in the same manner as increased temperature. Thus, hybridization conditions can be readily manipulated, and thus will generally be a method of choice depending on the desired results.

POLYNUCLEOTIDE IDENTIFICATION AND CHARACTERIZATION

Polynucleotides may be identified, prepared and/or manipulated using any of a variety of well established techniques. For example, a polynucleotide may be identified, as described in more detail below, by screening a microarray of cDNAs for tumor-associated expression (*i.e.*, expression that is at least two fold greater in a tumor than in normal tissue, as determined using a representative assay provided herein). Such screens may be performed, for example, using a Synteni microarray (Palo Alto, CA) according to the manufacturer's instructions (and essentially as described by Schena *et al.*, *Proc. Natl. Acad. Sci. USA* 93:10614-10619, 1996 and Heller *et al.*, *Proc. Natl. Acad. Sci. USA* 94:2150-2155, 1997). Alternatively, polynucleotides may be amplified from cDNA prepared from cells expressing the proteins described herein, such as prostate-specific cells. Such polynucleotides may be amplified via polymerase chain reaction (PCR). For this

approach, sequence-specific primers may be designed based on the sequences provided herein, and may be purchased or synthesized.

An amplified portion of a polynucleotide of the present invention may be used to isolate a full length gene from a suitable library (e.g., a prostate tumor cDNA library) using well known techniques. Within such techniques, a library (cDNA or genomic) is screened using one or more polynucleotide probes or primers suitable for amplification. Preferably, a library is size-selected to include larger molecules. Random primed libraries may also be preferred for identifying 5' and upstream regions of genes. Genomic libraries are preferred for obtaining introns and extending 5' sequences.

For hybridization techniques, a partial sequence may be labeled (e.g., by nick-translation or end-labeling with ^{32}P) using well known techniques. A bacterial or bacteriophage library is then generally screened by hybridizing filters containing denatured bacterial colonies (or lawns containing phage plaques) with the labeled probe (see Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989). Hybridizing colonies or plaques are selected and expanded, and the DNA is isolated for further analysis. cDNA clones may be analyzed to determine the amount of additional sequence by, for example, PCR using a primer from the partial sequence and a primer from the vector. Restriction maps and partial sequences may be generated to identify one or more overlapping clones. The complete sequence may then be determined using standard techniques, which may involve generating a series of deletion clones. The resulting overlapping sequences can then be assembled into a single contiguous sequence. A full length cDNA molecule can be generated by ligating suitable fragments, using well known techniques.

Alternatively, there are numerous amplification techniques for obtaining a full length coding sequence from a partial cDNA sequence. Within such techniques, amplification is generally performed via PCR. Any of a variety of commercially available kits may be used to perform the amplification step. Primers may be designed using, for example, software well known in the art. Primers are preferably 22-30 nucleotides in length, have a GC content of at least 50% and anneal to the target sequence at temperatures

of about 68°C to 72°C. The amplified region may be sequenced as described above, and overlapping sequences assembled into a contiguous sequence.

One such amplification technique is inverse PCR (*see Triglia et al., Nucl. Acids Res. 16:8186, 1988*), which uses restriction enzymes to generate a fragment in the
 5 known region of the gene. The fragment is then circularized by intramolecular ligation and used as a template for PCR with divergent primers derived from the known region. Within an alternative approach, sequences adjacent to a partial sequence may be retrieved by amplification with a primer to a linker sequence and a primer specific to a known region. The amplified sequences are typically subjected to a second round of amplification with the
 10 same linker primer and a second primer specific to the known region. A variation on this procedure, which employs two primers that initiate extension in opposite directions from the known sequence, is described in WO 96/38591. Another such technique is known as "rapid amplification of cDNA ends" or RACE. This technique involves the use of an internal primer and an external primer, which hybridizes to a polyA region or vector
 15 sequence, to identify sequences that are 5' and 3' of a known sequence. Additional techniques include capture PCR (*Lagerstrom et al., PCR Methods Applic. 1:111-19, 1991*) and walking PCR (*Parker et al., Nucl. Acids. Res. 19:3055-60, 1991*). Other methods employing amplification may also be employed to obtain a full length cDNA sequence.

In certain instances, it is possible to obtain a full length cDNA sequence by
 20 analysis of sequences provided in an expressed sequence tag (EST) database, such as that available from GenBank. Searches for overlapping ESTs may generally be performed using well known programs (*e.g., NCBI BLAST searches*), and such ESTs may be used to generate a contiguous full length sequence. Full length DNA sequences may also be obtained by analysis of genomic fragments.

25 POLYNUCLEOTIDE EXPRESSION IN HOST CELLS

In other embodiments of the invention, polynucleotide sequences or fragments thereof which encode polypeptides of the invention, or fusion proteins or functional equivalents thereof, may be used in recombinant DNA molecules to direct

expression of a polypeptide in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences that encode substantially the same or a functionally equivalent amino acid sequence may be produced and these sequences may be used to clone and express a given polypeptide.

5 As will be understood by those of skill in the art, it may be advantageous in some instances to produce polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce a recombinant RNA transcript having desirable properties, such as a half-life which is longer
10 than that of a transcript generated from the naturally occurring sequence.

 Moreover, the polynucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter polypeptide encoding sequences for a variety of reasons, including but not limited to, alterations which modify the cloning, processing, and/or expression of the gene product. For example, DNA
15 shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. In addition, site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so forth.

20 In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences may be ligated to a heterologous sequence to encode a fusion protein. For example, to screen peptide libraries for inhibitors of polypeptide activity, it may be useful to encode a chimeric protein that can be recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site
25 located between the polypeptide-encoding sequence and the heterologous protein sequence, so that the polypeptide may be cleaved and purified away from the heterologous moiety.

 Sequences encoding a desired polypeptide may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M. H. *et al.* (1980) *Nucl. Acids Res. Symp. Ser.* 215-223, Horn, T. *et al.* (1980) *Nucl. Acids Res. Symp. Ser.*

225-232). Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of a polypeptide, or a portion thereof. For example, peptide synthesis can be performed using various solid-phase techniques (Roberge, J. Y. *et al.* (1995) *Science* 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer, Palo Alto, CA).

A newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (*e.g.*, Creighton, T. (1983) *Proteins, Structures and Molecular Principles*, WH Freeman and Co., New York, N.Y.) or other comparable techniques available in the art. The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (*e.g.*, the Edman degradation procedure). Additionally, the amino acid sequence of a polypeptide, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from other proteins, or any part thereof, to produce a variant polypeptide.

In order to express a desired polypeptide, the nucleotide sequences encoding the polypeptide, or functional equivalents, may be inserted into appropriate expression vector, *i.e.*, a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding a polypeptide of interest and appropriate transcriptional and translational control elements. These methods include *in vitro* recombinant DNA techniques, synthetic techniques, and *in vivo* genetic recombination. Such techniques are described in Sambrook, J. *et al.* (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview, N.Y., and Ausubel, F. M. *et al.* (1989) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, N.Y.

A variety of expression vector/host systems may be utilized to contain and express polynucleotide sequences. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (*e.g.*, baculovirus); plant cell systems transformed

with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids); or animal cell systems.

The "control elements" or "regulatory sequences" present in an expression
 5 vector are those non-translated regions of the vector--enhancers, promoters, 5' and 3' untranslated regions--which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when
 10 cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the PBLUESCRIPT phagemid (Stratagene, La Jolla, Calif.) or PSPORT1 plasmid (Gibco BRL, Gaithersburg, MD) and the like may be used. In mammalian cell systems, promoters from mammalian genes or from mammalian viruses are generally preferred. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding a polypeptide,
 15 vectors based on SV40 or EBV may be advantageously used with an appropriate selectable marker.

In bacterial systems, a number of expression vectors may be selected depending upon the use intended for the expressed polypeptide. For example, when large quantities are needed, for example for the induction of antibodies, vectors which direct
 20 high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the sequence encoding the polypeptide of interest may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of .beta.-galactosidase so that a hybrid protein is produced;
 25 pIN vectors (Van Heeke, G. and S. M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509); and the like. pGEX Vectors (Promega, Madison, Wis.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins

made in such systems may be designed to include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, *Saccharomyces cerevisiae*, a number of vectors containing
 5 constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel *et al.* (supra) and Grant *et al.* (1987) *Methods Enzymol.* 153:516-544.

In cases where plant expression vectors are used, the expression of
 sequences encoding polypeptides may be driven by any of a number of promoters. For
 10 example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311. Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. *et al.* (1984) *EMBO J.* 3:1671-1680; Broglie, R. *et al.* (1984) *Science* 224:838-843; and Winter, J. *et al.* (1991) *Results Probl.*
 15 *Cell Differ.* 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, Hobbs, S. or Murry, L. E. in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, N.Y.; pp. 191-196).

20 An insect system may also be used to express a polypeptide of interest. For example, in one such system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in *Trichoplusia* larvae. The sequences encoding the polypeptide may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the
 25 polyhedrin promoter. Successful insertion of the polypeptide-encoding sequence will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein. The recombinant viruses may then be used to infect, for example, *S. frugiperda* cells or *Trichoplusia* larvae in which the polypeptide of interest may be expressed (Engelhard, E. K. *et al.* (1994) *Proc. Natl. Acad. Sci.* 91 :3224-3227).

In mammalian host cells, a number of viral-based expression systems are generally available. For example, in cases where an adenovirus is used as an expression vector, sequences encoding a polypeptide of interest may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing the polypeptide in infected host cells (Logan, J. and Shenk, T. (1984) *Proc. Natl. Acad. Sci.* 81:3655-3659). In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be used to achieve more efficient translation of sequences encoding a polypeptide of interest. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding the polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. *et al.* (1994) *Results Probl. Cell Differ.* 20:125-162).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells such as CHO, HeLa, MDCK, HEK293, and WI38, which have specific cellular machinery and characteristic mechanisms

for such post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is generally preferred. For example, cell lines which stably express a polynucleotide of interest may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. *et al.* (1977) *Cell* 11:223-32) and adenine phosphoribosyltransferase (Lowy, I. *et al.* (1990) *Cell* 22:817-23) genes which can be employed in tk.sup.- or aprt.sup.- cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. *et al.* (1980) *Proc. Natl. Acad. Sci.* 77:3567-70); npt, which confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin, F. *et al.* (1981) *J. Mol. Biol.* 150:1-14); and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, *supra*). Additional selectable genes have been described, for example, trpB, which allows cells to utilize indole in place of tryptophan, or hisD, which allows cells to utilize histinol in place of histidine (Hartman, S. C. and R. C. Mulligan (1988) *Proc. Natl. Acad. Sci.* 85:8047-51). Recently, the use of visible markers has gained popularity with such markers as anthocyanins, beta-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein

expression attributable to a specific vector system (Rhodes, C. A. *et al.* (1995) *Methods Mol. Biol.* 55:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding a polypeptide is inserted within a marker gene sequence, recombinant cells containing sequences can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a polypeptide-encoding sequence under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

Alternatively, host cells which contain and express a desired polynucleotide sequence may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

A variety of protocols for detecting and measuring the expression of polynucleotide-encoded products, using either polyclonal or monoclonal antibodies specific for the product are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on a given polypeptide may be preferred for some applications, but a competitive binding assay may also be employed. These and other assays are described, among other places, in Hampton, R. *et al.* (1990; *Serological Methods, a Laboratory Manual*, APS Press, St Paul, Minn.) and Maddox, D. E. *et al.* (1983; *J. Exp. Med.* 158:1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides include oligolabeling, nick translation, end-labeling or PCR amplification

using a labeled nucleotide. Alternatively, the sequences, or any portions thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled
 5 nucleotides. These procedures may be conducted using a variety of commercially available kits. Suitable reporter molecules or labels, which may be used include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with a polynucleotide sequence of interest may be
 10 cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides of the invention may be designed to contain signal sequences which direct secretion of the encoded polypeptide
 15 through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may be used to join sequences encoding a polypeptide of interest to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals,
 20 protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen, San Diego, Calif.) between the purification domain and the encoded polypeptide may be used to facilitate purification. One such expression vector
 25 provides for expression of a fusion protein containing a polypeptide of interest and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography) as described in Porath, J. *et al.* (1992, *Prot. Exp. Purif.* 3:263-281) while the enterokinase cleavage site provides a means for purifying the desired

polypeptide from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D. J. *et al.* (1993; *DNA Cell Biol.* 12:441-453).

In addition to recombinant production methods, polypeptides of the invention, and fragments thereof, may be produced by direct peptide synthesis using solid-phase techniques (Merrifield J. (1963) *J. Am. Chem. Soc.* 85:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Alternatively, various fragments may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

10 SITE-SPECIFIC MUTAGENESIS

Site-specific mutagenesis is a technique useful in the preparation of individual peptides, or biologically functional equivalent polypeptides, through specific mutagenesis of the underlying polynucleotides that encode them. The technique, well-known to those of skill in the art, further provides a ready ability to prepare and test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the DNA. Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Mutations may be employed in a selected polynucleotide sequence to improve, alter, decrease, modify, or otherwise change the properties of the polynucleotide itself, and/or alter the properties, activity, composition, stability, or primary sequence of the encoded polypeptide.

In certain embodiments of the present invention, the inventors contemplate the mutagenesis of the disclosed polynucleotide sequences to alter one or more properties of the encoded polypeptide, such as the antigenicity of a polypeptide vaccine. The techniques of site-specific mutagenesis are well-known in the art, and are widely used to

create variants of both polypeptides and polynucleotides. For example, site-specific mutagenesis is often used to alter a specific portion of a DNA molecule. In such embodiments, a primer comprising typically about 14 to about 25 nucleotides or so in length is employed, with about 5 to about 10 residues on both sides of the junction of the
 5 sequence being altered.

As will be appreciated by those of skill in the art, site-specific mutagenesis techniques have often employed a phage vector that exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage. These phage are readily commercially-available and their use is
 10 generally well-known to those skilled in the art. Double-stranded plasmids are also routinely employed in site directed mutagenesis that eliminates the step of transferring the gene of interest from a plasmid to a phage.

In general, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector or melting apart of two strands of a
 15 double-stranded vector that includes within its sequence a DNA sequence that encodes the desired peptide. An oligonucleotide primer bearing the desired mutated sequence is prepared, generally synthetically. This primer is then annealed with the single-stranded vector, and subjected to DNA polymerizing enzymes such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a
 20 heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is then used to transform appropriate cells, such as *E. coli* cells, and clones are selected which include recombinant vectors bearing the mutated sequence arrangement.

The preparation of sequence variants of the selected peptide-encoding DNA
 25 segments using site-directed mutagenesis provides a means of producing potentially useful species and is not meant to be limiting as there are other ways in which sequence variants of peptides and the DNA sequences encoding them may be obtained. For example, recombinant vectors encoding the desired peptide sequence may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence variants. Specific details regarding

these methods and protocols are found in the teachings of Maloy *et al.*, 1994; Segal, 1976; Prokop and Bajpai, 1991; Kuby, 1994; and Maniatis *et al.*, 1982, each incorporated herein by reference, for that purpose.

As used herein, the term “oligonucleotide directed mutagenesis procedure” refers to template-dependent processes and vector-mediated propagation which result in an increase in the concentration of a specific nucleic acid molecule relative to its initial concentration, or in an increase in the concentration of a detectable signal, such as amplification. As used herein, the term “oligonucleotide directed mutagenesis procedure” is intended to refer to a process that involves the template-dependent extension of a primer molecule. The term template dependent process refers to nucleic acid synthesis of an RNA or a DNA molecule wherein the sequence of the newly synthesized strand of nucleic acid is dictated by the well-known rules of complementary base pairing (see, for example, Watson, 1987). Typically, vector mediated methodologies involve the introduction of the nucleic acid fragment into a DNA or RNA vector, the clonal amplification of the vector, and the recovery of the amplified nucleic acid fragment. Examples of such methodologies are provided by U. S. Patent No. 4,237,224, specifically incorporated herein by reference in its entirety.

POLYNUCLEOTIDE AMPLIFICATION TECHNIQUES

A number of template dependent processes are available to amplify the target sequences of interest present in a sample. One of the best known amplification methods is the polymerase chain reaction (PCRTM) which is described in detail in U.S. Patent Nos. 4,683,195, 4,683,202 and 4,800,159, each of which is incorporated herein by reference in its entirety. Briefly, in PCRTM, two primer sequences are prepared which are complementary to regions on opposite complementary strands of the target sequence. An excess of deoxynucleoside triphosphates is added to a reaction mixture along with a DNA polymerase (*e.g.*, *Taq* polymerase). If the target sequence is present in a sample, the primers will bind to the target and the polymerase will cause the primers to be extended along the target sequence by adding on nucleotides. By raising and lowering the

temperature of the reaction mixture, the extended primers will dissociate from the target to form reaction products, excess primers will bind to the target and to the reaction product and the process is repeated. Preferably reverse transcription and PCR™ amplification procedure may be performed in order to quantify the amount of mRNA amplified.

5 Polymerase chain reaction methodologies are well known in the art.

Another method for amplification is the ligase chain reaction (referred to as LCR), disclosed in Eur. Pat. Appl. Publ. No. 320,308 (specifically incorporated herein by reference in its entirety). In LCR, two complementary probe pairs are prepared, and in the presence of the target sequence, each pair will bind to opposite complementary strands of the target such that they abut. In the presence of a ligase, the two probe pairs will link to form a single unit. By temperature cycling, as in PCR™, bound ligated units dissociate from the target and then serve as "target sequences" for ligation of excess probe pairs. U.S. Patent No. 4,883,750, incorporated herein by reference in its entirety, describes an alternative method of amplification similar to LCR for binding probe pairs to a target sequence.

Qbeta Replicase, described in PCT Intl. Pat. Appl. Publ. No. PCT/US87/00880, incorporated herein by reference in its entirety, may also be used as still another amplification method in the present invention. In this method, a replicative sequence of RNA that has a region complementary to that of a target is added to a sample in the presence of an RNA polymerase. The polymerase will copy the replicative sequence that can then be detected.

An isothermal amplification method, in which restriction endonucleases and ligases are used to achieve the amplification of target molecules that contain nucleotide 5'-[α-thio]triphosphates in one strand of a restriction site (Walker *et al.*, 1992, incorporated herein by reference in its entirety), may also be useful in the amplification of nucleic acids in the present invention.

Strand Displacement Amplification (SDA) is another method of carrying out isothermal amplification of nucleic acids which involves multiple rounds of strand displacement and synthesis, *i.e.* nick translation. A similar method, called Repair Chain

Reaction (RCR) is another method of amplification which may be useful in the present invention and is involves annealing several probes throughout a region targeted for amplification, followed by a repair reaction in which only two of the four bases are present. The other two bases can be added as biotinylated derivatives for easy detection. A similar approach is used in SDA.

Sequences can also be detected using a cyclic probe reaction (CPR). In CPR, a probe having a 3' and 5' sequences of non-target DNA and an internal or "middle" sequence of the target protein specific RNA is hybridized to DNA which is present in a sample. Upon hybridization, the reaction is treated with RNaseH, and the products of the probe are identified as distinctive products by generating a signal that is released after digestion. The original template is annealed to another cycling probe and the reaction is repeated. Thus, CPR involves amplifying a signal generated by hybridization of a probe to a target gene specific expressed nucleic acid.

Still other amplification methods described in Great Britain Pat. Appl. No. 2 02 328, and in PCT Intl. Pat. Appl. Publ. No. PCT/US89/01025, each of which is incorporated herein by reference in its entirety, may be used in accordance with the present invention. In the former application, "modified" primers are used in a PCR-like, template and enzyme dependent synthesis. The primers may be modified by labeling with a capture moiety (*e.g.*, biotin) and/or a detector moiety (*e.g.*, enzyme). In the latter application, an excess of labeled probes is added to a sample. In the presence of the target sequence, the probe binds and is cleaved catalytically. After cleavage, the target sequence is released intact to be bound by excess probe. Cleavage of the labeled probe signals the presence of the target sequence.

Other nucleic acid amplification procedures include transcription-based amplification systems (TAS) (Kwoh *et al.*, 1989; PCT Intl. Pat. Appl. Publ. No. WO 88/10315, incorporated herein by reference in its entirety), including nucleic acid sequence based amplification (NASBA) and 3SR. In NASBA, the nucleic acids can be prepared for amplification by standard phenol/chloroform extraction, heat denaturation of a sample, treatment with lysis buffer and minispin columns for isolation of DNA and RNA or

guanidinium chloride extraction of RNA. These amplification techniques involve annealing a primer that has sequences specific to the target sequence. Following polymerization, DNA/RNA hybrids are digested with RNase H while double stranded DNA molecules are heat-denatured again. In either case the single stranded DNA is made
5 fully double stranded by addition of second target-specific primer, followed by polymerization. The double stranded DNA molecules are then multiply transcribed by a polymerase such as T7 or SP6. In an isothermal cyclic reaction, the RNAs are reverse transcribed into DNA, and transcribed once again with a polymerase such as T7 or SP6. The resulting products, whether truncated or complete, indicate target-specific sequences.

10 Eur. Pat. Appl. Publ. No. 329,822, incorporated herein by reference in its entirety, disclose a nucleic acid amplification process involving cyclically synthesizing single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA), which may be used in accordance with the present invention. The ssRNA is a first template for a first primer oligonucleotide, which is elongated by reverse transcriptase (RNA-dependent DNA
15 polymerase). The RNA is then removed from resulting DNA:RNA duplex by the action of ribonuclease H (RNase H, an RNase specific for RNA in a duplex with either DNA or RNA). The resultant ssDNA is a second template for a second primer, which also includes the sequences of an RNA polymerase promoter (exemplified by T7 RNA polymerase) 5' to its homology to its template. This primer is then extended by DNA polymerase
20 (exemplified by the large "Klenow" fragment of *E. coli* DNA polymerase I), resulting as a double-stranded DNA ("dsDNA") molecule, having a sequence identical to that of the original RNA between the primers and having additionally, at one end, a promoter sequence. This promoter sequence can be used by the appropriate RNA polymerase to make many RNA copies of the DNA. These copies can then re-enter the cycle leading to
25 very swift amplification. With proper choice of enzymes, this amplification can be done isothermally without addition of enzymes at each cycle. Because of the cyclical nature of this process, the starting sequence can be chosen to be in the form of either DNA or RNA.

PCT Intl. Pat. Appl. Publ. No. WO 89/06700, incorporated herein by reference in its entirety, disclose a nucleic acid sequence amplification scheme based on the

hybridization of a promoter/primer sequence to a target single-stranded DNA ("ssDNA") followed by transcription of many RNA copies of the sequence. This scheme is not cyclic; *i.e.* new templates are not produced from the resultant RNA transcripts. Other amplification methods include "RACE" (Frohman, 1990), and "one-sided PCR" (Ohara, 5 1989) which are well-known to those of skill in the art.

Methods based on ligation of two (or more) oligonucleotides in the presence of nucleic acid having the sequence of the resulting "di-oligonucleotide", thereby amplifying the di-oligonucleotide (Wu and Dean, 1996, incorporated herein by reference in its entirety), may also be used in the amplification of DNA sequences of the present 10 invention.

BIOLOGICAL FUNCTIONAL EQUIVALENTS

Modification and changes may be made in the structure of the polynucleotides and polypeptides of the present invention and still obtain a functional molecule that encodes a polypeptide with desirable characteristics. As mentioned above, it 15 is often desirable to introduce one or more mutations into a specific polynucleotide sequence. In certain circumstances, the resulting encoded polypeptide sequence is altered by this mutation, or in other cases, the sequence of the polypeptide is unchanged by one or more mutations in the encoding polynucleotide.

When it is desirable to alter the amino acid sequence of a polypeptide to 20 create an equivalent, or even an improved, second-generation molecule, the amino acid changes may be achieved by changing one or more of the codons of the encoding DNA sequence, according to Table 1.

For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with 25 structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence, and, of course, its underlying DNA coding sequence, and

nevertheless obtain a protein with like properties. It is thus contemplated by the inventors that various changes may be made in the peptide sequences of the disclosed compositions, or corresponding DNA sequences which encode said peptides without appreciable loss of their biological utility or activity.

5

TABLE 1

Amino Acids			Codons						
Alanine	Ala	A	GCA	GCC	GCG	GCU			
Cysteine	Cys	C	UGC	UGU					
Aspartic acid	Asp	D	GAC	GAU					
Glutamic acid	Glu	E	GAA	GAG					
Phenylalanine	Phe	F	UUC	UUU					
Glycine	Gly	G	GGA	GGC	GGG	GGU			
Histidine	His	H	CAC	CAU					
Isoleucine	Ile	I	AUA	AUC	AUU				
Lysine	Lys	K	AAA	AAG					
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU	
Methionine	Met	M	AUG						
Asparagine	Asn	N	AAC	AAU					
Proline	Pro	P	CCA	CCC	CCG	CCU			
Glutamine	Gln	Q	CAA	CAG					
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU	
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU	
Threonine	Thr	T	ACA	ACC	ACG	ACU			
Valine	Val	V	GUA	GUC	GUG	GUU			
Tryptophan	Trp	W	UGG						
Tyrosine	Tyr	Y	UAC	UAU					

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive

biologic function on a protein is generally understood in the art (Kyte and Doolittle, 1982, incorporated herein by reference). It is accepted that the relative hydropathic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like. Each amino acid has been assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics (Kyte and Doolittle, 1982). These values are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (−0.4); threonine (−0.7); serine (−0.8); tryptophan (−0.9); tyrosine (−1.3); proline (−1.6); histidine (−3.2); glutamate (−3.5); glutamine (−3.5); aspartate (−3.5); asparagine (−3.5); lysine (−3.9); and arginine (−4.5).

It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydropathic index or score and still result in a protein with similar biological activity, *i.e.* still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose hydropathic indices are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred. It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U. S. Patent 4,554,101 (specifically incorporated herein by reference in its entirety), states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

As detailed in U. S. Patent 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 \pm 1); glutamate (+3.0 \pm 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (−0.4); proline (−0.5 \pm 1); alanine (−0.5); histidine (−0.5); cysteine (−1.0); methionine (−1.3); valine (−1.5); leucine (−1.8); isoleucine (−1.8); tyrosine (−2.3); phenylalanine (−2.5); tryptophan (−3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent protein. In such changes, the

substitution of amino acids whose hydrophilicity values are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

As outlined above, amino acid substitutions are generally therefore based on the relative similarity of the amino acid side-chain substituents, for example, their hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

In addition, any polynucleotide may be further modified to increase stability *in vivo*. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends; the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages in the backbone; and/or the inclusion of nontraditional bases such as inosine, queosine and wybutosine, as well as acetyl- methyl-, thio- and other modified forms of adenine, cytidine, guanine, thymine and uridine.

IN VIVO POLYNUCLEOTIDE DELIVERY TECHNIQUES

In additional embodiments, genetic constructs comprising one or more of the polynucleotides of the invention are introduced into cells *in vivo*. This may be achieved using any of a variety of well known approaches, several of which are outlined below for the purpose of illustration.

1. ADENOVIRUS

One of the preferred methods for *in vivo* delivery of one or more nucleic acid sequences involves the use of an adenovirus expression vector. "Adenovirus expression vector" is meant to include those constructs containing adenovirus sequences sufficient to (a) support packaging of the construct and (b) to express a polynucleotide that has been cloned therein in a sense or antisense orientation. Of course, in the context of an antisense construct, expression does not require that the gene product be synthesized.

The expression vector comprises a genetically engineered form of an adenovirus. Knowledge of the genetic organization of adenovirus, a 36 kb, linear, double-stranded DNA virus, allows substitution of large pieces of adenoviral DNA with foreign sequences up to 7 kb (Grunhaus and Horwitz, 1992). In contrast to retrovirus, the
 5 adenoviral infection of host cells does not result in chromosomal integration because adenoviral DNA can replicate in an episomal manner without potential genotoxicity. Also, adenoviruses are structurally stable, and no genome rearrangement has been detected after extensive amplification. Adenovirus can infect virtually all epithelial cells regardless of their cell cycle stage. So far, adenoviral infection appears to be linked only to mild disease
 10 such as acute respiratory disease in humans.

Adenovirus is particularly suitable for use as a gene transfer vector because of its mid-sized genome, ease of manipulation, high titer, wide target-cell range and high infectivity. Both ends of the viral genome contain 100-200 base pair inverted repeats (ITRs), which are *cis* elements necessary for viral DNA replication and packaging. The
 15 early (E) and late (L) regions of the genome contain different transcription units that are divided by the onset of viral DNA replication. The E1 region (E1A and E1B) encodes proteins responsible for the regulation of transcription of the viral genome and a few cellular genes. The expression of the E2 region (E2A and E2B) results in the synthesis of the proteins for viral DNA replication. These proteins are involved in DNA replication,
 20 late gene expression and host cell shut-off (Renan, 1990). The products of the late genes, including the majority of the viral capsid proteins, are expressed only after significant processing of a single primary transcript issued by the major late promoter (MLP). The MLP, (located at 16.8 m.u.) is particularly efficient during the late phase of infection, and all the mRNA's issued from this promoter possess a 5'-tripartite leader (TPL) sequence
 25 which makes them preferred mRNA's for translation.

In a current system, recombinant adenovirus is generated from homologous recombination between shuttle vector and provirus vector. Due to the possible recombination between two proviral vectors, wild-type adenovirus may be generated from

this process. Therefore, it is critical to isolate a single clone of virus from an individual plaque and examine its genomic structure.

Generation and propagation of the current adenovirus vectors, which are replication deficient, depend on a unique helper cell line, designated 293, which was transformed from human embryonic kidney cells by Ad5 DNA fragments and constitutively expresses E1 proteins (Graham *et al.*, 1977). Since the E3 region is dispensable from the adenovirus genome (Jones and Shenk, 1978), the current adenovirus vectors, with the help of 293 cells, carry foreign DNA in either the E1, the D3 or both regions (Graham and Prevec, 1991). In nature, adenovirus can package approximately 105% of the wild-type genome (Ghosh-Choudhury *et al.*, 1987), providing capacity for about 2 extra kB of DNA. Combined with the approximately 5.5 kB of DNA that is replaceable in the E1 and E3 regions, the maximum capacity of the current adenovirus vector is under 7.5 kB, or about 15% of the total length of the vector. More than 80% of the adenovirus viral genome remains in the vector backbone and is the source of vector-borne cytotoxicity. Also, the replication deficiency of the E1-deleted virus is incomplete. For example, leakage of viral gene expression has been observed with the currently available vectors at high multiplicities of infection (MOI) (Mulligan, 1993).

Helper cell lines may be derived from human cells such as human embryonic kidney cells, muscle cells, hematopoietic cells or other human embryonic mesenchymal or epithelial cells. Alternatively, the helper cells may be derived from the cells of other mammalian species that are permissive for human adenovirus. Such cells include, *e.g.*, Vero cells or other monkey embryonic mesenchymal or epithelial cells. As stated above, the currently preferred helper cell line is 293.

Recently, Racher *et al.* (1995) disclosed improved methods for culturing 293 cells and propagating adenovirus. In one format, natural cell aggregates are grown by inoculating individual cells into 1 liter siliconized spinner flasks (Techne, Cambridge, UK) containing 100-200 ml of medium. Following stirring at 40 rpm, the cell viability is estimated with trypan blue. In another format, Fibra-Cel microcarriers (Bibby Sterlin, Stone, UK) (5 g/l) is employed as follows. A cell inoculum, resuspended in 5 ml of

medium, is added to the carrier (50 ml) in a 250 ml Erlenmeyer flask and left stationary, with occasional agitation, for 1 to 4 h. The medium is then replaced with 50 ml of fresh medium and shaking initiated. For virus production, cells are allowed to grow to about 80% confluence, after which time the medium is replaced (to 25% of the final volume) and
 5 adenovirus added at an MOI of 0.05. Cultures are left stationary overnight, following which the volume is increased to 100% and shaking commenced for another 72 h.

Other than the requirement that the adenovirus vector be replication defective, or at least conditionally defective, the nature of the adenovirus vector is not believed to be crucial to the successful practice of the invention. The adenovirus may be of
 10 any of the 42 different known serotypes or subgroups A-F. Adenovirus type 5 of subgroup C is the preferred starting material in order to obtain a conditional replication-defective adenovirus vector for use in the present invention, since Adenovirus type 5 is a human adenovirus about which a great deal of biochemical and genetic information is known, and it has historically been used for most constructions employing adenovirus as a vector.

15 As stated above, the typical vector according to the present invention is replication defective and will not have an adenovirus E1 region. Thus, it will be most convenient to introduce the polynucleotide encoding the gene of interest at the position from which the E1-coding sequences have been removed. However, the position of insertion of the construct within the adenovirus sequences is not critical to the invention.
 20 The polynucleotide encoding the gene of interest may also be inserted in lieu of the deleted E3 region in E3 replacement vectors as described by Karlsson *et al.* (1986) or in the E4 region where a helper cell line or helper virus complements the E4 defect.

Adenovirus is easy to grow and manipulate and exhibits broad host range *in vitro* and *in vivo*. This group of viruses can be obtained in high titers, *e.g.*, 10^9 - 10^{11} plaque-
 25 forming units per ml, and they are highly infective. The life cycle of adenovirus does not require integration into the host cell genome. The foreign genes delivered by adenovirus vectors are episomal and, therefore, have low genotoxicity to host cells. No side effects have been reported in studies of vaccination with wild-type adenovirus (Couch *et al.*, 1963;

Top *et al.*, 1971), demonstrating their safety and therapeutic potential as *in vivo* gene transfer vectors.

Adenovirus vectors have been used in eukaryotic gene expression (Levrero *et al.*, 1991; Gomez-Foix *et al.*, 1992) and vaccine development (Grunhaus and Horwitz, 1992; Graham and Prevec, 1992). Recently, animal studies suggested that recombinant adenovirus could be used for gene therapy (Stratford-Perricaudet and Perricaudet, 1991; Stratford-Perricaudet *et al.*, 1990; Rich *et al.*, 1993). Studies in administering recombinant adenovirus to different tissues include trachea instillation (Rosenfeld *et al.*, 1991; Rosenfeld *et al.*, 1992), muscle injection (Ragot *et al.*, 1993), peripheral intravenous injections (Herz and Gerard, 1993) and stereotactic inoculation into the brain (Le Gal La Salle *et al.*, 1993).

2. RETROVIRUSES

The retroviruses are a group of single-stranded RNA viruses characterized by an ability to convert their RNA to double-stranded DNA in infected cells by a process of reverse-transcription (Coffin, 1990). The resulting DNA then stably integrates into cellular chromosomes as a provirus and directs synthesis of viral proteins. The integration results in the retention of the viral gene sequences in the recipient cell and its descendants. The retroviral genome contains three genes, gag, pol, and env that code for capsid proteins, polymerase enzyme, and envelope components, respectively. A sequence found upstream from the gag gene contains a signal for packaging of the genome into virions. Two long terminal repeat (LTR) sequences are present at the 5' and 3' ends of the viral genome. These contain strong promoter and enhancer sequences and are also required for integration in the host cell genome (Coffin, 1990).

In order to construct a retroviral vector, a nucleic acid encoding one or more oligonucleotide or polynucleotide sequences of interest is inserted into the viral genome in the place of certain viral sequences to produce a virus that is replication-defective. In order to produce virions, a packaging cell line containing the gag, pol, and env genes but without the LTR and packaging components is constructed (Mann *et al.*, 1983). When a

recombinant plasmid containing a cDNA, together with the retroviral LTR and packaging sequences is introduced into this cell line (by calcium phosphate precipitation for example), the packaging sequence allows the RNA transcript of the recombinant plasmid to be packaged into viral particles, which are then secreted into the culture media (Nicolas and Rubenstein, 1988; Temin, 1986; Mann *et al.*, 1983). The media containing the recombinant retroviruses is then collected, optionally concentrated, and used for gene transfer. Retroviral vectors are able to infect a broad variety of cell types. However, integration and stable expression require the division of host cells (Paskind *et al.*, 1975).

A novel approach designed to allow specific targeting of retrovirus vectors was recently developed based on the chemical modification of a retrovirus by the chemical addition of lactose residues to the viral envelope. This modification could permit the specific infection of hepatocytes *via* sialoglycoprotein receptors.

A different approach to targeting of recombinant retroviruses was designed in which biotinylated antibodies against a retroviral envelope protein and against a specific cell receptor were used. The antibodies were coupled *via* the biotin components by using streptavidin (Roux *et al.*, 1989). Using antibodies against major histocompatibility complex class I and class II antigens, they demonstrated the infection of a variety of human cells that bore those surface antigens with an ecotropic virus *in vitro* (Roux *et al.*, 1989).

3. ADENO-ASSOCIATED VIRUSES

AAV (Ridgeway, 1988; Hermonat and Muzyczka, 1984) is a parovirus, discovered as a contamination of adenoviral stocks. It is a ubiquitous virus (antibodies are present in 85% of the US human population) that has not been linked to any disease. It is also classified as a dependovirus, because its replications is dependent on the presence of a helper virus, such as adenovirus. Five serotypes have been isolated, of which AAV-2 is the best characterized. AAV has a single-stranded linear DNA that is encapsidated into capsid proteins VP1, VP2 and VP3 to form an icosahedral virion of 20 to 24 nm in diameter (Muzyczka and McLaughlin, 1988).

The AAV DNA is approximately 4.7 kilobases long. It contains two open reading frames and is flanked by two ITRs. There are two major genes in the AAV genome: *rep* and *cap*. The *rep* gene codes for proteins responsible for viral replications, whereas *cap* codes for capsid protein VP1-3. Each ITR forms a T-shaped hairpin structure. These terminal repeats are the only essential *cis* components of the AAV for chromosomal integration. Therefore, the AAV can be used as a vector with all viral coding sequences removed and replaced by the cassette of genes for delivery. Three viral promoters have been identified and named p5, p19, and p40, according to their map position. Transcription from p5 and p19 results in production of rep proteins, and transcription from p40 produces the capsid proteins (Hermonat and Muzyczka, 1984).

There are several factors that prompted researchers to study the possibility of using rAAV as an expression vector. One is that the requirements for delivering a gene to integrate into the host chromosome are surprisingly few. It is necessary to have the 145-bp ITRs, which are only 6% of the AAV genome. This leaves room in the vector to assemble a 4.5-kb DNA insertion. While this carrying capacity may prevent the AAV from delivering large genes, it is amply suited for delivering the antisense constructs of the present invention.

AAV is also a good choice of delivery vehicles due to its safety. There is a relatively complicated rescue mechanism: not only wild type adenovirus but also AAV genes are required to mobilize rAAV. Likewise, AAV is not pathogenic and not associated with any disease. The removal of viral coding sequences minimizes immune reactions to viral gene expression, and therefore, rAAV does not evoke an inflammatory response.

4. OTHER VIRAL VECTORS AS EXPRESSION CONSTRUCTS

Other viral vectors may be employed as expression constructs in the present invention for the delivery of oligonucleotide or polynucleotide sequences to a host cell. Vectors derived from viruses such as vaccinia virus (Ridgeway, 1988; Coupar *et al.*, 1988), lentiviruses, polio viruses and herpes viruses may be employed. They offer several

attractive features for various mammalian cells (Friedmann, 1989; Ridgeway, 1988; Coupar *et al.*, 1988; Horwich *et al.*, 1990).

With the recent recognition of defective hepatitis B viruses, new insight was gained into the structure-function relationship of different viral sequences. *In vitro* studies showed that the virus could retain the ability for helper-dependent packaging and reverse transcription despite the deletion of up to 80% of its genome (Horwich *et al.*, 1990). This suggested that large portions of the genome could be replaced with foreign genetic material. The hepatotropism and persistence (integration) were particularly attractive properties for liver-directed gene transfer. Chang *et al.* (1991) introduced the chloramphenicol acetyltransferase (CAT) gene into duck hepatitis B virus genome in the place of the polymerase, surface, and pre-surface coding sequences. It was cotransfected with wild-type virus into an avian hepatoma cell line. Culture media containing high titers of the recombinant virus were used to infect primary duckling hepatocytes. Stable CAT gene expression was detected for at least 24 days after transfection (Chang *et al.*, 1991).

5. NON-VIRAL VECTORS

In order to effect expression of the oligonucleotide or polynucleotide sequences of the present invention, the expression construct must be delivered into a cell. This delivery may be accomplished *in vitro*, as in laboratory procedures for transforming cells lines, or *in vivo* or *ex vivo*, as in the treatment of certain disease states. As described above, one preferred mechanism for delivery is *via* viral infection where the expression construct is encapsulated in an infectious viral particle.

Once the expression construct has been delivered into the cell the nucleic acid encoding the desired oligonucleotide or polynucleotide sequences may be positioned and expressed at different sites. In certain embodiments, the nucleic acid encoding the construct may be stably integrated into the genome of the cell. This integration may be in the specific location and orientation *via* homologous recombination (gene replacement) or it may be integrated in a random, non-specific location (gene augmentation). In yet further embodiments, the nucleic acid may be stably maintained in the cell as a separate, episomal

segment of DNA. Such nucleic acid segments or "episomes" encode sequences sufficient to permit maintenance and replication independent of or in synchronization with the host cell cycle. How the expression construct is delivered to a cell and where in the cell the nucleic acid remains is dependent on the type of expression construct employed.

5 In certain embodiments of the invention, the expression construct comprising one or more oligonucleotide or polynucleotide sequences may simply consist of naked recombinant DNA or plasmids. Transfer of the construct may be performed by any of the methods mentioned above which physically or chemically permeabilize the cell membrane. This is particularly applicable for transfer *in vitro* but it may be applied to *in*
 10 *vivo* use as well. Dubensky *et al.* (1984) successfully injected polyomavirus DNA in the form of calcium phosphate precipitates into liver and spleen of adult and newborn mice demonstrating active viral replication and acute infection. Benvenisty and Reshef (1986) also demonstrated that direct intraperitoneal injection of calcium phosphate-precipitated plasmids results in expression of the transfected genes. It is envisioned that DNA encoding
 15 a gene of interest may also be transferred in a similar manner *in vivo* and express the gene product.

 Another embodiment of the invention for transferring a naked DNA expression construct into cells may involve particle bombardment. This method depends on the ability to accelerate DNA-coated microprojectiles to a high velocity allowing them
 20 to pierce cell membranes and enter cells without killing them (Klein *et al.*, 1987). Several devices for accelerating small particles have been developed. One such device relies on a high voltage discharge to generate an electrical current, which in turn provides the motive force (Yang *et al.*, 1990). The microprojectiles used have consisted of biologically inert substances such as tungsten or gold beads.

25 Selected organs including the liver, skin, and muscle tissue of rats and mice have been bombarded *in vivo* (Yang *et al.*, 1990; Zelenin *et al.*, 1991). This may require surgical exposure of the tissue or cells, to eliminate any intervening tissue between the gun and the target organ, *i.e.* *ex vivo* treatment. Again, DNA encoding a particular gene may be delivered *via* this method and still be incorporated by the present invention.

ANTISENSE OLIGONUCLEOTIDES

The end result of the flow of genetic information is the synthesis of protein. DNA is transcribed by polymerases into messenger RNA and translated on the ribosome to yield a folded, functional protein. Thus there are several steps along the route where protein synthesis can be inhibited. The native DNA segment coding for a polypeptide described herein, as all such mammalian DNA strands, has two strands: a sense strand and an antisense strand held together by hydrogen bonding. The messenger RNA coding for polypeptide has the same nucleotide sequence as the sense DNA strand except that the DNA thymidine is replaced by uridine. Thus, synthetic antisense nucleotide sequences will bind to a mRNA and inhibit expression of the protein encoded by that mRNA.

The targeting of antisense oligonucleotides to mRNA is thus one mechanism to shut down protein synthesis, and, consequently, represents a powerful and targeted therapeutic approach. For example, the synthesis of polygalacturonase and the muscarine type 2 acetylcholine receptor are inhibited by antisense oligonucleotides directed to their respective mRNA sequences (U. S. Patent 5,739,119 and U. S. Patent 5,759,829, each specifically incorporated herein by reference in its entirety). Further, examples of antisense inhibition have been demonstrated with the nuclear protein cyclin, the multiple drug resistance gene (MDG1), ICAM-1, E-selectin, STK-1, striatal GABA_A receptor and human EGF (Jaskulski *et al.*, 1988; Vasanthakumar and Ahmed, 1989; Peris *et al.*, 1998; U. S. Patent 5,801,154; U. S. Patent 5,789,573; U. S. Patent 5,718,709 and U. S. Patent 5,610,288, each specifically incorporated herein by reference in its entirety). Antisense constructs have also been described that inhibit and can be used to treat a variety of abnormal cellular proliferations, *e.g.* cancer (U. S. Patent 5,747,470; U. S. Patent 5,591,317 and U. S. Patent 5,783,683, each specifically incorporated herein by reference in its entirety).

Therefore, in exemplary embodiments, the invention provides oligonucleotide sequences that comprise all, or a portion of, any sequence that is capable of specifically binding to polynucleotide sequence described herein, or a complement thereof. In one embodiment, the antisense oligonucleotides comprise DNA or derivatives thereof.

In another embodiment, the oligonucleotides comprise RNA or derivatives thereof. In a third embodiment, the oligonucleotides are modified DNAs comprising a phosphorothioated modified backbone. In a fourth embodiment, the oligonucleotide sequences comprise peptide nucleic acids or derivatives thereof. In each case, preferred
 5 compositions comprise a sequence region that is complementary, and more preferably substantially-complementary, and even more preferably, completely complementary to one or more portions of polynucleotides disclosed herein.

Selection of antisense compositions specific for a given gene sequence is based upon analysis of the chosen target sequence (*i.e.* in these illustrative examples the rat
 10 and human sequences) and determination of secondary structure, T_m , binding energy, relative stability, and antisense compositions were selected based upon their relative inability to form dimers, hairpins, or other secondary structures that would reduce or prohibit specific binding to the target mRNA in a host cell.

Highly preferred target regions of the mRNA, are those which are at or near
 15 the AUG translation initiation codon, and those sequences which were substantially complementary to 5' regions of the mRNA. These secondary structure analyses and target site selection considerations were performed using v.4 of the OLIGO primer analysis software (Rychlik, 1997) and the BLASTN 2.0.5 algorithm software (Altschul *et al.*, 1997).

The use of an antisense delivery method employing a short peptide vector,
 20 termed MPG (27 residues), is also contemplated. The MPG peptide contains a hydrophobic domain derived from the fusion sequence of HIV gp41 and a hydrophilic domain from the nuclear localization sequence of SV40 T-antigen (Morris *et al.*, 1997). It has been demonstrated that several molecules of the MPG peptide coat the antisense oligonucleotides and can be delivered into cultured mammalian cells in less than 1 hour
 25 with relatively high efficiency (90%). Further, the interaction with MPG strongly increases both the stability of the oligonucleotide to nuclease and the ability to cross the plasma membrane (Morris *et al.*, 1997).

RIBOZYMES

Although proteins traditionally have been used for catalysis of nucleic acids, another class of macromolecules has emerged as useful in this endeavor. Ribozymes are RNA-protein complexes that cleave nucleic acids in a site-specific fashion. Ribozymes have specific catalytic domains that possess endonuclease activity (Kim and Cech, 1987; Gerlach *et al.*, 1987; Forster and Symons, 1987). For example, a large number of ribozymes accelerate phosphoester transfer reactions with a high degree of specificity, often cleaving only one of several phosphoesters in an oligonucleotide substrate (Cech *et al.*, 1981; Michel and Westhof, 1990; Reinhold-Hurek and Shub, 1992). This specificity has been attributed to the requirement that the substrate bind via specific base-pairing interactions to the internal guide sequence ("IGS") of the ribozyme prior to chemical reaction.

Ribozyme catalysis has primarily been observed as part of sequence-specific cleavage/ligation reactions involving nucleic acids (Joyce, 1989; Cech *et al.*, 1981). For example, U. S. Patent No. 5,354,855 (specifically incorporated herein by reference) reports that certain ribozymes can act as endonucleases with a sequence specificity greater than that of known ribonucleases and approaching that of the DNA restriction enzymes. Thus, sequence-specific ribozyme-mediated inhibition of gene expression may be particularly suited to therapeutic applications (Scanlon *et al.*, 1991; Sarver *et al.*, 1990). Recently, it was reported that ribozymes elicited genetic changes in some cells lines to which they were applied; the altered genes included the oncogenes H-*ras*, c-*fos* and genes of HIV. Most of this work involved the modification of a target mRNA, based on a specific mutant codon that is cleaved by a specific ribozyme.

Six basic varieties of naturally-occurring enzymatic RNAs are known presently. Each can catalyze the hydrolysis of RNA phosphodiester bonds *in trans* (and thus can cleave other RNA molecules) under physiological conditions. In general, enzymatic nucleic acids act by first binding to a target RNA. Such binding occurs through the target binding portion of a enzymatic nucleic acid which is held in close proximity to an enzymatic portion of the molecule that acts to cleave the target RNA. Thus, the enzymatic

nucleic acid first recognizes and then binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cut the target RNA. Strategic cleavage of such a target RNA will destroy its ability to direct synthesis of an encoded protein. After an enzymatic nucleic acid has bound and cleaved its RNA target, it
 5 is released from that RNA to search for another target and can repeatedly bind and cleave new targets.

The enzymatic nature of a ribozyme is advantageous over many technologies, such as antisense technology (where a nucleic acid molecule simply binds to a nucleic acid target to block its translation) since the concentration of ribozyme necessary
 10 to affect a therapeutic treatment is lower than that of an antisense oligonucleotide. This advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single ribozyme molecule is able to cleave many molecules of target RNA. In addition, the ribozyme is a highly specific inhibitor, with the specificity of inhibition depending not only on the base pairing mechanism of binding to the target RNA, but also on the mechanism of
 15 target RNA cleavage. Single mismatches, or base-substitutions, near the site of cleavage can completely eliminate catalytic activity of a ribozyme. Similar mismatches in antisense molecules do not prevent their action (Woolf *et al.*, 1992). Thus, the specificity of action of a ribozyme is greater than that of an antisense oligonucleotide binding the same RNA site.

20 The enzymatic nucleic acid molecule may be formed in a hammerhead, hairpin, a hepatitis δ virus, group I intron or RNaseP RNA (in association with an RNA guide sequence) or Neurospora VS RNA motif. Examples of hammerhead motifs are described by Rossi *et al.* (1992). Examples of hairpin motifs are described by Hampel
 25 *et al.* (Eur. Pat. Appl. Publ. No. EP 0360257), Hampel and Tritz (1989), Hampel *et al.* (1990) and U. S. Patent 5,631,359 (specifically incorporated herein by reference). An example of the hepatitis δ virus motif is described by Perrotta and Been (1992); an example of the RNaseP motif is described by Guerrier-Takada *et al.* (1983); Neurospora VS RNA ribozyme motif is described by Collins (Saville and Collins, 1990; Saville and Collins, 1991; Collins and Olive, 1993); and an example of the Group I intron is described in (U. S.

Patent 4,987,071, specifically incorporated herein by reference). All that is important in an enzymatic nucleic acid molecule of this invention is that it has a specific substrate binding site which is complementary to one or more of the target gene RNA regions, and that it have nucleotide sequences within or surrounding that substrate binding site which impart an RNA cleaving activity to the molecule. Thus the ribozyme constructs need not be limited to specific motifs mentioned herein.

In certain embodiments, it may be important to produce enzymatic cleaving agents which exhibit a high degree of specificity for the RNA of a desired target, such as one of the sequences disclosed herein. The enzymatic nucleic acid molecule is preferably targeted to a highly conserved sequence region of a target mRNA. Such enzymatic nucleic acid molecules can be delivered exogenously to specific cells as required. Alternatively, the ribozymes can be expressed from DNA or RNA vectors that are delivered to specific cells.

Small enzymatic nucleic acid motifs (*e.g.*, of the hammerhead or the hairpin structure) may also be used for exogenous delivery. The simple structure of these molecules increases the ability of the enzymatic nucleic acid to invade targeted regions of the mRNA structure. Alternatively, catalytic RNA molecules can be expressed within cells from eukaryotic promoters (*e.g.*, Scanlon *et al.*, 1991; Kashani-Sabet *et al.*, 1992; Dropulic *et al.*, 1992; Weerasinghe *et al.*, 1991; Ojwang *et al.*, 1992; Chen *et al.*, 1992; Sarver *et al.*, 1990). Those skilled in the art realize that any ribozyme can be expressed in eukaryotic cells from the appropriate DNA vector. The activity of such ribozymes can be augmented by their release from the primary transcript by a second ribozyme (Int. Pat. Appl. Publ. No. WO 93/23569, and Int. Pat. Appl. Publ. No. WO 94/02595, both hereby incorporated by reference; Ohkawa *et al.*, 1992; Taira *et al.*, 1991; and Ventura *et al.*, 1993).

Ribozymes may be added directly, or can be complexed with cationic lipids, lipid complexes, packaged within liposomes, or otherwise delivered to target cells. The RNA or RNA complexes can be locally administered to relevant tissues *ex vivo*, or *in vivo* through injection, aerosol inhalation, infusion pump or stent, with or without their incorporation in biopolymers.

Ribozymes may be designed as described in Int. Pat. Appl. Publ. No. WO 93/23569 and Int. Pat. Appl. Publ. No. WO 94/02595, each specifically incorporated herein by reference) and synthesized to be tested *in vitro* and *in vivo*, as described. Such ribozymes can also be optimized for delivery. While specific examples are provided, those
 5 in the art will recognize that equivalent RNA targets in other species can be utilized when necessary.

Hammerhead or hairpin ribozymes may be individually analyzed by computer folding (Jaeger *et al.*, 1989) to assess whether the ribozyme sequences fold into the appropriate secondary structure. Those ribozymes with unfavorable intramolecular
 10 interactions between the binding arms and the catalytic core are eliminated from consideration. Varying binding arm lengths can be chosen to optimize activity. Generally, at least 5 or so bases on each arm are able to bind to, or otherwise interact with, the target RNA.

Ribozymes of the hammerhead or hairpin motif may be designed to anneal
 15 to various sites in the mRNA message, and can be chemically synthesized. The method of synthesis used follows the procedure for normal RNA synthesis as described in Usman *et al.* (1987) and in Scaringe *et al.* (1990) and makes use of common nucleic acid protecting and coupling groups, such as dimethoxytrityl at the 5'-end, and phosphoramidites at the 3'-end. Average stepwise coupling yields are typically >98%.
 20 Hairpin ribozymes may be synthesized in two parts and annealed to reconstruct an active ribozyme (Chowrira and Burke, 1992). Ribozymes may be modified extensively to enhance stability by modification with nuclease resistant groups, for example, 2'-amino, 2'-C-allyl, 2'-fluoro, 2'-o-methyl, 2'-H (for a review see *e.g.*, Usman and Cedergren, 1992). Ribozymes may be purified by gel electrophoresis using general methods or by high
 25 pressure liquid chromatography and resuspended in water.

Ribozyme activity can be optimized by altering the length of the ribozyme binding arms, or chemically synthesizing ribozymes with modifications that prevent their degradation by serum ribonucleases (see *e.g.*, Int. Pat. Appl. Publ. No. WO 92/07065; Perrault *et al.*, 1990; Pieken *et al.*, 1991; Usman and Cedergren, 1992; Int. Pat. Appl. Publ.

No. WO 93/15187; Int. Pat. Appl. Publ. No. WO 91/03162; Eur. Pat. Appl. Publ. No. 92110298.4; U. S. Patent 5,334,711; and Int. Pat. Appl. Publ. No. WO 94/13688, which describe various chemical modifications that can be made to the sugar moieties of enzymatic RNA molecules), modifications which enhance their efficacy in cells, and
 5 removal of stem II bases to shorten RNA synthesis times and reduce chemical requirements.

Sullivan *et al.* (Int. Pat. Appl. Publ. No. WO 94/02595) describes the general methods for delivery of enzymatic RNA molecules. Ribozymes may be administered to cells by a variety of methods known to those familiar to the art, including,
 10 but not restricted to, encapsulation in liposomes, by iontophoresis, or by incorporation into other vehicles, such as hydrogels, cyclodextrins, biodegradable nanocapsules, and bioadhesive microspheres. For some indications, ribozymes may be directly delivered *ex vivo* to cells or tissues with or without the aforementioned vehicles. Alternatively, the RNA/vehicle combination may be locally delivered by direct inhalation, by direct injection
 15 or by use of a catheter, infusion pump or stent. Other routes of delivery include, but are not limited to, intravascular, intramuscular, subcutaneous or joint injection, aerosol inhalation, oral (tablet or pill form), topical, systemic, ocular, intraperitoneal and/or intrathecal delivery. More detailed descriptions of ribozyme delivery and administration are provided in Int. Pat. Appl. Publ. No. WO 94/02595 and Int. Pat. Appl. Publ. No. WO 93/23569, each
 20 specifically incorporated herein by reference.

Another means of accumulating high concentrations of a ribozyme(s) within cells is to incorporate the ribozyme-encoding sequences into a DNA expression vector. Transcription of the ribozyme sequences are driven from a promoter for eukaryotic RNA polymerase I (pol I), RNA polymerase II (pol II), or RNA polymerase III (pol III).
 25 Transcripts from pol II or pol III promoters will be expressed at high levels in all cells; the levels of a given pol II promoter in a given cell type will depend on the nature of the gene regulatory sequences (enhancers, silencers, *etc.*) present nearby. Prokaryotic RNA polymerase promoters may also be used, providing that the prokaryotic RNA polymerase enzyme is expressed in the appropriate cells (Elroy-Stein and Moss, 1990; Gao and Huang,

1993; Lieber *et al.*, 1993; Zhou *et al.*, 1990). Ribozymes expressed from such promoters can function in mammalian cells (*e.g.* Kashani-Saber *et al.*, 1992; Ojwang *et al.*, 1992; Chen *et al.*, 1992; Yu *et al.*, 1993; L'Huillier *et al.*, 1992; Lisiewicz *et al.*, 1993). Such transcription units can be incorporated into a variety of vectors for introduction into
 5 mammalian cells, including but not restricted to, plasmid DNA vectors, viral DNA vectors (such as adenovirus or adeno-associated vectors), or viral RNA vectors (such as retroviral, semliki forest virus, sindbis virus vectors).

Ribozymes may be used as diagnostic tools to examine genetic drift and mutations within diseased cells. They can also be used to assess levels of the target RNA
 10 molecule. The close relationship between ribozyme activity and the structure of the target RNA allows the detection of mutations in any region of the molecule which alters the base-pairing and three-dimensional structure of the target RNA. By using multiple ribozymes, one may map nucleotide changes which are important to RNA structure and function *in vitro*, as well as in cells and tissues. Cleavage of target RNAs with ribozymes may be used
 15 to inhibit gene expression and define the role (essentially) of specified gene products in the progression of disease. In this manner, other genetic targets may be defined as important mediators of the disease. These studies will lead to better treatment of the disease progression by affording the possibility of combinational therapies (*e.g.*, multiple ribozymes targeted to different genes, ribozymes coupled with known small molecule
 20 inhibitors, or intermittent treatment with combinations of ribozymes and/or other chemical or biological molecules). Other *in vitro* uses of ribozymes are well known in the art, and include detection of the presence of mRNA associated with an IL-5 related condition. Such RNA is detected by determining the presence of a cleavage product after treatment with a ribozyme using standard methodology.

25 **PEPTIDE NUCLEIC ACIDS**

In certain embodiments, the inventors contemplate the use of peptide nucleic acids (PNAs) in the practice of the methods of the invention. PNA is a DNA mimic in which the nucleobases are attached to a pseudopeptide backbone (Good and

Nielsen, 1997). PNA is able to be utilized in a number methods that traditionally have used RNA or DNA. Often PNA sequences perform better in techniques than the corresponding RNA or DNA sequences and have utilities that are not inherent to RNA or DNA. A review of PNA including methods of making, characteristics of, and methods of using, is provided
 5 by Corey (1997) and is incorporated herein by reference. As such, in certain embodiments, one may prepare PNA sequences that are complementary to one or more portions of the ACE mRNA sequence, and such PNA compositions may be used to regulate, alter, decrease, or reduce the translation of ACE-specific mRNA, and thereby alter the level of ACE activity in a host cell to which such PNA compositions have been administered.

10 PNAs have 2-aminoethyl-glycine linkages replacing the normal phosphodiester backbone of DNA (Nielsen *et al.*, 1991; Hanvey *et al.*, 1992; Hyrup and Nielsen, 1996; Neilsen, 1996). This chemistry has three important consequences: firstly, in contrast to DNA or phosphorothioate oligonucleotides, PNAs are neutral molecules; secondly, PNAs are achiral, which avoids the need to develop a stereoselective synthesis;
 15 and thirdly, PNA synthesis uses standard Boc (Dueholm *et al.*, 1994) or Fmoc (Thomson *et al.*, 1995) protocols for solid-phase peptide synthesis, although other methods, including a modified Merrifield method, have been used (Christensen *et al.*, 1995).

PNA monomers or ready-made oligomers are commercially available from PerSeptive Biosystems (Framingham, MA). PNA syntheses by either Boc or Fmoc
 20 protocols are straightforward using manual or automated protocols (Norton *et al.*, 1995). The manual protocol lends itself to the production of chemically modified PNAs or the simultaneous synthesis of families of closely related PNAs.

As with peptide synthesis, the success of a particular PNA synthesis will depend on the properties of the chosen sequence. For example, while in theory PNAs can
 25 incorporate any combination of nucleotide bases, the presence of adjacent purines can lead to deletions of one or more residues in the product. In expectation of this difficulty, it is suggested that, in producing PNAs with adjacent purines, one should repeat the coupling of residues likely to be added inefficiently. This should be followed by the purification of PNAs by reverse-phase high-pressure liquid chromatography (Norton *et al.*, 1995)

providing yields and purity of product similar to those observed during the synthesis of peptides.

Modifications of PNAs for a given application may be accomplished by coupling amino acids during solid-phase synthesis or by attaching compounds that contain a carboxylic acid group to the exposed N-terminal amine. Alternatively, PNAs can be modified after synthesis by coupling to an introduced lysine or cysteine. The ease with which PNAs can be modified facilitates optimization for better solubility or for specific functional requirements. Once synthesized, the identity of PNAs and their derivatives can be confirmed by mass spectrometry. Several studies have made and utilized modifications of PNAs (Norton *et al.*, 1995; Haaima *et al.*, 1996; Stetsenko *et al.*, 1996; Petersen *et al.*, 1995; Ulmann *et al.*, 1996; Koch *et al.*, 1995; Orum *et al.*, 1995; Footer *et al.*, 1996; Griffith *et al.*, 1995; Kremsky *et al.*, 1996; Pardridge *et al.*, 1995; Boffa *et al.*, 1995; Landsdorp *et al.*, 1996; Gambacorti-Passerini *et al.*, 1996; Armitage *et al.*, 1997; Seeger *et al.*, 1997; Ruskowski *et al.*, 1997). U.S. Patent No. 5,700,922 discusses PNA-DNA-PNA chimeric molecules and their uses in diagnostics, modulating protein in organisms, and treatment of conditions susceptible to therapeutics.

In contrast to DNA and RNA, which contain negatively charged linkages, the PNA backbone is neutral. In spite of this dramatic alteration, PNAs recognize complementary DNA and RNA by Watson-Crick pairing (Egholm *et al.*, 1993), validating the initial modeling by Nielsen *et al.* (1991). PNAs lack 3' to 5' polarity and can bind in either parallel or antiparallel fashion, with the antiparallel mode being preferred (Egholm *et al.*, 1993).

Hybridization of DNA oligonucleotides to DNA and RNA is destabilized by electrostatic repulsion between the negatively charged phosphate backbones of the complementary strands. By contrast, the absence of charge repulsion in PNA-DNA or PNA-RNA duplexes increases the melting temperature (T_m) and reduces the dependence of T_m on the concentration of mono- or divalent cations (Nielsen *et al.*, 1991). The enhanced rate and affinity of hybridization are significant because they are responsible for the surprising ability of PNAs to perform strand invasion of complementary sequences within

relaxed double-stranded DNA. In addition, the efficient hybridization at inverted repeats suggests that PNAs can recognize secondary structure effectively within double-stranded DNA. Enhanced recognition also occurs with PNAs immobilized on surfaces, and Wang *et al.* have shown that support-bound PNAs can be used to detect hybridization events (Wang
5 *et al.*, 1996).

One might expect that tight binding of PNAs to complementary sequences would also increase binding to similar (but not identical) sequences, reducing the sequence specificity of PNA recognition. As with DNA hybridization, however, selective recognition can be achieved by balancing oligomer length and incubation temperature.
10 Moreover, selective hybridization of PNAs is encouraged by PNA-DNA hybridization being less tolerant of base mismatches than DNA-DNA hybridization. For example, a single mismatch within a 16 bp PNA-DNA duplex can reduce the T_m by up to 15°C (Egholm *et al.*, 1993). This high level of discrimination has allowed the development of several PNA-based strategies for the analysis of point mutations (Wang *et al.*, 1996;
15 Carlsson *et al.*, 1996; Thiede *et al.*, 1996; Webb and Hurskainen, 1996; Perry-O'Keefe *et al.*, 1996).

High-affinity binding provides clear advantages for molecular recognition and the development of new applications for PNAs. For example, 11-13 nucleotide PNAs inhibit the activity of telomerase, a ribonucleo-protein that extends telomere ends using an
20 essential RNA template, while the analogous DNA oligomers do not (Norton *et al.*, 1996).

Neutral PNAs are more hydrophobic than analogous DNA oligomers, and this can lead to difficulty solubilizing them at neutral pH, especially if the PNAs have a high purine content or if they have the potential to form secondary structures. Their solubility can be enhanced by attaching one or more positive charges to the PNA termini
25 (Nielsen *et al.*, 1991).

Findings by Allfrey and colleagues suggest that strand invasion will occur spontaneously at sequences within chromosomal DNA (Boffa *et al.*, 1995; Boffa *et al.*, 1996). These studies targeted PNAs to triplet repeats of the nucleotides CAG and used this recognition to purify transcriptionally active DNA (Boffa *et al.*, 1995) and to inhibit

transcription (Boffa *et al.*, 1996). This result suggests that if PNAs can be delivered within cells then they will have the potential to be general sequence-specific regulators of gene expression. Studies and reviews concerning the use of PNAs as antisense and anti-gene agents include Nielsen *et al.* (1993b), Hanvey *et al.* (1992), and Good and Nielsen (1997).
 5 Koppelhus *et al.* (1997) have used PNAs to inhibit HIV-1 inverse transcription, showing that PNAs may be used for antiviral therapies.

Methods of characterizing the antisense binding properties of PNAs are discussed in Rose (1993) and Jensen *et al.* (1997). Rose uses capillary gel electrophoresis to determine binding of PNAs to their complementary oligonucleotide, measuring the
 10 relative binding kinetics and stoichiometry. Similar types of measurements were made by Jensen *et al.* using BIAcore™ technology.

Other applications of PNAs include use in DNA strand invasion (Nielsen *et al.*, 1991), antisense inhibition (Hanvey *et al.*, 1992), mutational analysis (Orum *et al.*, 1993), enhancers of transcription (Mollegaard *et al.*, 1994), nucleic acid purification (Orum
 15 *et al.*, 1995), isolation of transcriptionally active genes (Boffa *et al.*, 1995), blocking of transcription factor binding (Vickers *et al.*, 1995), genome cleavage (Veselkov *et al.*, 1996), biosensors (Wang *et al.*, 1996), *in situ* hybridization (Thisted *et al.*, 1996), and in a alternative to Southern blotting (Perry-O'Keefe, 1996).

POLYPEPTIDE COMPOSITIONS

20 The present invention, in other aspects, provides polypeptide compositions. Generally, a polypeptide of the invention will be an isolated polypeptide (or an epitope, variant, or active fragment thereof) derived from a mammalian species. Preferably, the polypeptide is encoded by a polynucleotide sequence disclosed herein or a sequence which hybridizes under moderately stringent conditions to a polynucleotide sequence disclosed
 25 herein. Alternatively, the polypeptide may be defined as a polypeptide which comprises a contiguous amino acid sequence from an amino acid sequence disclosed herein, or which polypeptide comprises an entire amino acid sequence disclosed herein.

In the present invention, a polypeptide composition is also understood to comprise one or more polypeptides that are immunologically reactive with antibodies generated against a polypeptide of the invention, particularly a polypeptide having the amino acid sequence disclosed in SEQ ID NO: 112-114, 172, 176, 178, 327, 329, 331, 336, 339, 376-380, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534, 537-551, 553-568, 573-586, 588-590, 592, 706-708, 775, 776, 778 and 780, or active fragments, variants or biological functional equivalents thereof.

Likewise, a polypeptide composition of the present invention is understood to comprise one or more polypeptides that are capable of eliciting antibodies that are immunologically reactive with one or more polypeptides encoded by one or more contiguous nucleic acid sequences contained in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824, or to active fragments, or to variants thereof, or to one or more nucleic acid sequences which hybridize to one or more of these sequences under conditions of moderate to high stringency. Particularly illustrative polypeptides include the amino acid sequence disclosed in SEQ ID NO: 112-114, 172, 176, 178, 327, 329, 331, 336, 339, 376-380, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534, 537-551, 553-568, 573-586, 588-590, 592, 706-708, 775, 776, 778 and 780.

As used herein, an active fragment of a polypeptide includes a whole or a portion of a polypeptide which is modified by conventional techniques, *e.g.*, mutagenesis, or by addition, deletion, or substitution, but which active fragment exhibits substantially the same structure function, antigenicity, etc., as a polypeptide as described herein.

In certain illustrative embodiments, the polypeptides of the invention will comprise at least an immunogenic portion of a prostate-specific protein or a variant thereof, as described herein. As noted above, a "prostate-specific protein" is a protein that is expressed by prostate cells. Proteins that are prostate-specific proteins also react detectably within an immunoassay (such as an ELISA) with antisera from a patient with prostate cancer. Polypeptides as described herein may be of any length. Additional sequences

derived from the native protein and/or heterologous sequences may be present, and such sequences may (but need not) possess further immunogenic or antigenic properties.

An "immunogenic portion," as used herein is a portion of a protein that is recognized (*i.e.*, specifically bound) by a B-cell and/or T-cell surface antigen receptor.

- 5 Such immunogenic portions generally comprise at least 5 amino acid residues, more preferably at least 10, and still more preferably at least 20 amino acid residues of a prostate-specific protein or a variant thereof. Certain preferred immunogenic portions include peptides in which an N-terminal leader sequence and/or transmembrane domain have been deleted. Other preferred immunogenic portions may contain a small N- and/or
- 10 C-terminal deletion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids), relative to the mature protein.

- Immunogenic portions may generally be identified using well known techniques, such as those summarized in Paul, *Fundamental Immunology*, 3rd ed., 243-247 (Raven Press, 1993) and references cited therein. Such techniques include screening
- 15 polypeptides for the ability to react with antigen-specific antibodies, antisera and/or T-cell lines or clones. As used herein, antisera and antibodies are "antigen-specific" if they specifically bind to an antigen (*i.e.*, they react with the protein in an ELISA or other immunoassay, and do not react detectably with unrelated proteins). Such antisera and antibodies may be prepared as described herein, and using well known techniques. An
- 20 immunogenic portion of a native prostate-specific protein is a portion that reacts with such antisera and/or T-cells at a level that is not substantially less than the reactivity of the full length polypeptide (*e.g.*, in an ELISA and/or T-cell reactivity assay). Such immunogenic portions may react within such assays at a level that is similar to or greater than the reactivity of the full length polypeptide. Such screens may generally be performed using
- 25 methods well known to those of ordinary skill in the art, such as those described in Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. For example, a polypeptide may be immobilized on a solid support and contacted with patient sera to allow binding of antibodies within the sera to the immobilized polypeptide.

Unbound sera may then be removed and bound antibodies detected using, for example, ¹²⁵I-labeled Protein A.

As noted above, a composition may comprise a variant of a native prostate-specific protein. A polypeptide "variant," as used herein, is a polypeptide that differs from a native prostate-specific protein in one or more substitutions, deletions, additions and/or insertions, such that the immunogenicity of the polypeptide is not substantially diminished. In other words, the ability of a variant to react with antigen-specific antisera may be enhanced or unchanged, relative to the native protein, or may be diminished by less than 50%, and preferably less than 20%, relative to the native protein. Such variants may generally be identified by modifying one of the above polypeptide sequences and evaluating the reactivity of the modified polypeptide with antigen-specific antibodies or antisera as described herein. Preferred variants include those in which one or more portions, such as an N-terminal leader sequence or transmembrane domain, have been removed. Other preferred variants include variants in which a small portion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids) has been removed from the N- and/or C-terminal of the mature protein.

Polypeptide variants encompassed by the present invention include those exhibiting at least about 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or more identity (determined as described above) to the polypeptides disclosed herein.

Preferably, a variant contains conservative substitutions. A "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydropathic nature of the polypeptide to be substantially unchanged. Amino acid substitutions may generally be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic nature of the residues. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine,

isoleucine and valine; glycine and alanine; asparagine and glutamine; and serine, threonine, phenylalanine and tyrosine. Other groups of amino acids that may represent conservative changes include: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his. A variant may also, or alternatively, contain nonconservative changes. In a preferred embodiment, variant polypeptides differ from a native sequence by substitution, deletion or addition of five amino acids or fewer. Variants may also (or alternatively) be modified by, for example, the deletion or addition of amino acids that have minimal influence on the immunogenicity, secondary structure and hydropathic nature of the polypeptide.

As noted above, polypeptides may comprise a signal (or leader) sequence at the N-terminal end of the protein, which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification or identification of the polypeptide (*e.g.*, poly-His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

Polypeptides may be prepared using any of a variety of well known techniques. Recombinant polypeptides encoded by DNA sequences as described above may be readily prepared from the DNA sequences using any of a variety of expression vectors known to those of ordinary skill in the art. Expression may be achieved in any appropriate host cell that has been transformed or transfected with an expression vector containing a DNA molecule that encodes a recombinant polypeptide. Suitable host cells include prokaryotes, yeast, and higher eukaryotic cells, such as mammalian cells and plant cells. Preferably, the host cells employed are *E. coli*, yeast or a mammalian cell line such as COS or CHO. Supernatants from suitable host/vector systems which secrete recombinant protein or polypeptide into culture media may be first concentrated using a commercially available filter. Following concentration, the concentrate may be applied to a suitable purification matrix such as an affinity matrix or an ion exchange resin. Finally, one or more reverse phase HPLC steps can be employed to further purify a recombinant polypeptide.

Portions and other variants having less than about 100 amino acids, and generally less than about 50 amino acids, may also be generated by synthetic means, using techniques well known to those of ordinary skill in the art. For example, such polypeptides may be synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. *See Merrifield, J. Am. Chem. Soc. 85:2149-2146, 1963.* Equipment for automated synthesis of polypeptides is commercially available from suppliers such as Perkin Elmer/Applied BioSystems Division (Foster City, CA), and may be operated according to the manufacturer's instructions.

Within certain specific embodiments, a polypeptide may be a fusion protein that comprises multiple polypeptides as described herein, or that comprises at least one polypeptide as described herein and an unrelated sequence, such as a known tumor protein. A fusion partner may, for example, assist in providing T helper epitopes (an immunological fusion partner), preferably T helper epitopes recognized by humans, or may assist in expressing the protein (an expression enhancer) at higher yields than the native recombinant protein. Certain preferred fusion partners are both immunological and expression enhancing fusion partners. Other fusion partners may be selected so as to increase the solubility of the protein or to enable the protein to be targeted to desired intracellular compartments. Still further fusion partners include affinity tags, which facilitate purification of the protein.

Fusion proteins may generally be prepared using standard techniques, including chemical conjugation. Preferably, a fusion protein is expressed as a recombinant protein, allowing the production of increased levels, relative to a non-fused protein, in an expression system. Briefly, DNA sequences encoding the polypeptide components may be assembled separately, and ligated into an appropriate expression vector. The 3' end of the DNA sequence encoding one polypeptide component is ligated, with or without a peptide linker, to the 5' end of a DNA sequence encoding the second polypeptide component so that the reading frames of the sequences are in phase. This permits translation into a single fusion protein that retains the biological activity of both component polypeptides.

A peptide linker sequence may be employed to separate the first and second polypeptide components by a distance sufficient to ensure that each polypeptide folds into its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion protein using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea *et al.*, *Gene* 40:39-46, 1985; Murphy *et al.*, *Proc. Natl. Acad. Sci. USA* 83:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. The linker sequence may generally be from 1 to about 50 amino acids in length. Linker sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

The ligated DNA sequences are operably linked to suitable transcriptional or translational regulatory elements. The regulatory elements responsible for expression of DNA are located only 5' to the DNA sequence encoding the first polypeptides. Similarly, stop codons required to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

Fusion proteins are also provided. Such proteins comprise a polypeptide as described herein together with an unrelated immunogenic protein. Preferably the immunogenic protein is capable of eliciting a recall response. Examples of such proteins include tetanus, tuberculosis and hepatitis proteins (*see*, for example, Stoute *et al.* *New Engl. J. Med.*, 336:86-91, 1997).

Within preferred embodiments, an immunological fusion partner is derived from protein D, a surface protein of the gram-negative bacterium *Haemophilus influenza B* (WO 91/18926). Preferably, a protein D derivative comprises approximately the first third

of the protein (e.g., the first N-terminal 100-110 amino acids), and a protein D derivative may be lipidated. Within certain preferred embodiments, the first 109 residues of a Lipoprotein D fusion partner is included on the N-terminus to provide the polypeptide with additional exogenous T-cell epitopes and to increase the expression level in *E. coli* (thus
 5 functioning as an expression enhancer). The lipid tail ensures optimal presentation of the antigen to antigen presenting cells. Other fusion partners include the non-structural protein from influenzae virus, NS1 (hemagglutinin). Typically, the N-terminal 81 amino acids are used, although different fragments that include T-helper epitopes may be used.

In another embodiment, the immunological fusion partner is the protein
 10 known as LYTA, or a portion thereof (preferably a C-terminal portion). LYTA is derived from *Streptococcus pneumoniae*, which synthesizes an N-acetyl-L-alanine amidase known as amidase LYTA (encoded by the *LytA* gene; *Gene* 43:265-292, 1986). LYTA is an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LYTA protein is responsible for the affinity to the choline or to
 15 some choline analogues such as DEAE. This property has been exploited for the development of *E. coli* C-LYTA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LYTA fragment at the amino terminus has been described (see *Biotechnology* 10:795-798, 1992). Within a preferred embodiment, a repeat portion of LYTA may be incorporated into a fusion protein. A
 20 repeat portion is found in the C-terminal region starting at residue 178. A particularly preferred repeat portion incorporates residues 188-305.

In general, polypeptides (including fusion proteins) and polynucleotides as described herein are isolated. An "isolated" polypeptide or polynucleotide is one that is removed from its original environment. For example, a naturally-occurring protein is
 25 isolated if it is separated from some or all of the coexisting materials in the natural system. Preferably, such polypeptides are at least about 90% pure, more preferably at least about 95% pure and most preferably at least about 99% pure. A polynucleotide is considered to be isolated if, for example, it is cloned into a vector that is not a part of the natural environment.

BINDING AGENTS

The present invention further provides agents, such as antibodies and antigen-binding fragments thereof, that specifically bind to a prostate-specific protein. As used herein, an antibody, or antigen-binding fragment thereof, is said to "specifically bind" to a prostate-specific protein if it reacts at a detectable level (within, for example, an ELISA) with a prostate-specific protein, and does not react detectably with unrelated proteins under similar conditions. As used herein, "binding" refers to a noncovalent association between two separate molecules such that a complex is formed. The ability to bind may be evaluated by, for example, determining a binding constant for the formation of the complex. The binding constant is the value obtained when the concentration of the complex is divided by the product of the component concentrations. In general, two compounds are said to "bind," in the context of the present invention, when the binding constant for complex formation exceeds about 10^3 L/mol. The binding constant may be determined using methods well known in the art.

Binding agents may be further capable of differentiating between patients with and without a cancer, such as prostate cancer, using the representative assays provided herein. In other words, antibodies or other binding agents that bind to a prostate-specific protein will generate a signal indicating the presence of a cancer in at least about 20% of patients with the disease, and will generate a negative signal indicating the absence of the disease in at least about 90% of individuals without the cancer. To determine whether a binding agent satisfies this requirement, biological samples (*e.g.*, blood, sera, sputum, urine and/or tumor biopsies) from patients with and without a cancer (as determined using standard clinical tests) may be assayed as described herein for the presence of polypeptides that bind to the binding agent. It will be apparent that a statistically significant number of samples with and without the disease should be assayed. Each binding agent should satisfy the above criteria; however, those of ordinary skill in the art will recognize that binding agents may be used in combination to improve sensitivity.

Any agent that satisfies the above requirements may be a binding agent. For example, a binding agent may be a ribosome, with or without a peptide component, an

RNA molecule or a polypeptide. In a preferred embodiment, a binding agent is an antibody or an antigen-binding fragment thereof. Antibodies may be prepared by any of a variety of techniques known to those of ordinary skill in the art. *See, e.g.,* Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In

5 general, antibodies can be produced by cell culture techniques, including the generation of monoclonal antibodies as described herein, or via transfection of antibody genes into suitable bacterial or mammalian cell hosts, in order to allow for the production of recombinant antibodies. In one technique, an immunogen comprising the polypeptide is initially injected into any of a wide variety of mammals (*e.g.,* mice, rats, rabbits, sheep or

10 goats). In this step, the polypeptides of this invention may serve as the immunogen without modification. Alternatively, particularly for relatively short polypeptides, a superior immune response may be elicited if the polypeptide is joined to a carrier protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen is injected into the animal host, preferably according to a predetermined schedule incorporating one or more

15 booster immunizations, and the animals are bled periodically. Polyclonal antibodies specific for the polypeptide may then be purified from such antisera by, for example, affinity chromatography using the polypeptide coupled to a suitable solid support.

Monoclonal antibodies specific for an antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein, *Eur. J. Immunol.*

20 6:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the desired specificity (*i.e.,* reactivity with the polypeptide of interest). Such cell lines may be produced, for example, from spleen cells obtained from an animal immunized as described above. The spleen cells are then immortalized by, for example, fusion with a myeloma cell

25 fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells and myeloma cells may be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, thymidine) selection.

After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and their culture supernatants tested for binding activity against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of growing hybridoma colonies. In addition, various techniques may be employed to enhance the yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from the ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and extraction. The polypeptides of this invention may be used in the purification process in, for example, an affinity chromatography step.

Within certain embodiments, the use of antigen-binding fragments of antibodies may be preferred. Such fragments include Fab fragments, which may be prepared using standard techniques. Briefly, immunoglobulins may be purified from rabbit serum by affinity chromatography on Protein A bead columns (Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988) and digested by papain to yield Fab and Fc fragments. The Fab and Fc fragments may be separated by affinity chromatography on protein A bead columns.

Monoclonal antibodies of the present invention may be coupled to one or more therapeutic agents. Suitable agents in this regard include radionuclides, differentiation inducers, drugs, toxins, and derivatives thereof. Preferred radionuclides include ^{90}Y , ^{123}I , ^{125}I , ^{131}I , ^{186}Re , ^{188}Re , ^{211}At , and ^{212}Bi . Preferred drugs include methotrexate, and pyrimidine and purine analogs. Preferred differentiation inducers include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, diphtheria toxin, cholera toxin, gelonin, *Pseudomonas* exotoxin, *Shigella* toxin, and pokeweed antiviral protein.

A therapeutic agent may be coupled (e.g., covalently bonded) to a suitable monoclonal antibody either directly or indirectly (e.g., via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a substituent

capable of reacting with the other. For example, a nucleophilic group, such as an amino or sulfhydryl group, on one may be capable of reacting with a carbonyl-containing group, such as an anhydride or an acid halide, or with an alkyl group containing a good leaving group (*e.g.*, a halide) on the other.

5 Alternatively, it may be desirable to couple a therapeutic agent and an antibody via a linker group. A linker group can function as a spacer to distance an antibody from an agent in order to avoid interference with binding capabilities. A linker group can also serve to increase the chemical reactivity of a substituent on an agent or an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity
10 may also facilitate the use of agents, or functional groups on agents, which otherwise would not be possible.

It will be evident to those skilled in the art that a variety of bifunctional or polyfunctional reagents, both homo- and hetero-functional (such as those described in the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group.
15 Coupling may be effected, for example, through amino groups, carboxyl groups, sulfhydryl groups or oxidized carbohydrate residues. There are numerous references describing such methodology, *e.g.*, U.S. Patent No. 4,671,958, to Rodwell *et al.*

Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a
20 linker group which is cleavable during or upon internalization into a cell. A number of different cleavable linker groups have been described. The mechanisms for the intracellular release of an agent from these linker groups include cleavage by reduction of a disulfide bond (*e.g.*, U.S. Patent No. 4,489,710, to Spitler), by irradiation of a photolabile bond (*e.g.*, U.S. Patent No. 4,625,014, to Senter *et al.*), by hydrolysis of derivatized amino
25 acid side chains (*e.g.*, U.S. Patent No. 4,638,045, to Kohn *et al.*), by serum complement-mediated hydrolysis (*e.g.*, U.S. Patent No. 4,671,958, to Rodwell *et al.*), and acid-catalyzed hydrolysis (*e.g.*, U.S. Patent No. 4,569,789, to Blattler *et al.*).

It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In

another embodiment, more than one type of agent may be coupled to one antibody. Regardless of the particular embodiment, immunoconjugates with more than one agent may be prepared in a variety of ways. For example, more than one agent may be coupled directly to an antibody molecule, or linkers that provide multiple sites for attachment can
 5 be used. Alternatively, a carrier can be used.

A carrier may bear the agents in a variety of ways, including covalent bonding either directly or via a linker group. Suitable carriers include proteins such as albumins (*e.g.*, U.S. Patent No. 4,507,234, to Kato *et al.*), peptides and polysaccharides such as aminodextran (*e.g.*, U.S. Patent No. 4,699,784, to Shih *et al.*). A carrier may also
 10 bear an agent by noncovalent bonding or by encapsulation, such as within a liposome vesicle (*e.g.*, U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for radionuclide agents include radiohalogenated small molecules and chelating compounds. For example, U.S. Patent No. 4,735,792 discloses representative radiohalogenated small molecules and their synthesis. A radionuclide chelate may be formed from chelating compounds that
 15 include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For example, U.S. Patent No. 4,673,562, to Davison *et al.* discloses representative chelating compounds and their synthesis.

A variety of routes of administration for the antibodies and immunoconjugates may be used. Typically, administration will be intravenous,
 20 intramuscular, subcutaneous or in the bed of a resected tumor. It will be evident that the precise dose of the antibody/immunoconjugate will vary depending upon the antibody used, the antigen density on the tumor, and the rate of clearance of the antibody.

T CELLS

Immunotherapeutic compositions may also, or alternatively, comprise T
 25 cells specific for a prostate-specific protein. Such cells may generally be prepared *in vitro* or *ex vivo*, using standard procedures. For example, T cells may be isolated from bone marrow, peripheral blood, or a fraction of bone marrow or peripheral blood of a patient, using a commercially available cell separation system, such as the Isolex™ System,

available from Nexell Therapeutics, Inc. (Irvine, CA; see also U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243). Alternatively, T cells may be derived from related or unrelated humans, non-human mammals, cell lines or cultures.

5 T cells may be stimulated with a prostate-specific polypeptide, polynucleotide encoding a prostate-specific polypeptide and/or an antigen presenting cell (APC) that expresses such a polypeptide. Such stimulation is performed under conditions and for a time sufficient to permit the generation of T cells that are specific for the polypeptide. Preferably, a prostate-specific polypeptide or polynucleotide is present within
10 a delivery vehicle, such as a microsphere, to facilitate the generation of specific T cells.

T cells are considered to be specific for a prostate-specific polypeptide if the T cells specifically proliferate, secrete cytokines or kill target cells coated with the polypeptide or expressing a gene encoding the polypeptide. T cell specificity may be evaluated using any of a variety of standard techniques. For example, within a chromium
15 release assay or proliferation assay, a stimulation index of more than two fold increase in lysis and/or proliferation, compared to negative controls, indicates T cell specificity. Such assays may be performed, for example, as described in Chen *et al.*, *Cancer Res.* 54:1065-1070, 1994. Alternatively, detection of the proliferation of T cells may be accomplished by a variety of known techniques. For example, T cell proliferation can be detected by
20 measuring an increased rate of DNA synthesis (*e.g.*, by pulse-labeling cultures of T cells with tritiated thymidine and measuring the amount of tritiated thymidine incorporated into DNA). Contact with a prostate-specific polypeptide (100 ng/ml - 100 µg/ml, preferably 200 ng/ml - 25 µg/ml) for 3 - 7 days should result in at least a two fold increase in proliferation of the T cells. Contact as described above for 2-3 hours should result in
25 activation of the T cells, as measured using standard cytokine assays in which a two fold increase in the level of cytokine release (*e.g.*, TNF or IFN-γ) is indicative of T cell activation (*see* Coligan *et al.*, *Current Protocols in Immunology*, vol. 1, Wiley Interscience (Greene 1998)). T cells that have been activated in response to a prostate-specific polypeptide, polynucleotide or polypeptide-expressing APC may be CD4⁺ and/or CD8⁺.

prostate-specific protein-specific T cells may be expanded using standard techniques. Within preferred embodiments, the T cells are derived from a patient, a related donor or an unrelated donor, and are administered to the patient following stimulation and expansion.

For therapeutic purposes, CD4⁺ or CD8⁺ T cells that proliferate in response to a prostate-specific polypeptide, polynucleotide or APC can be expanded in number either *in vitro* or *in vivo*. Proliferation of such T cells *in vitro* may be accomplished in a variety of ways. For example, the T cells can be re-exposed to a prostate-specific polypeptide, or a short peptide corresponding to an immunogenic portion of such a polypeptide, with or without the addition of T cell growth factors, such as interleukin-2, and/or stimulator cells that synthesize a prostate-specific polypeptide. Alternatively, one or more T cells that proliferate in the presence of a prostate-specific protein can be expanded in number by cloning. Methods for cloning cells are well known in the art, and include limiting dilution.

PHARMACEUTICAL COMPOSITIONS

In additional embodiments, the present invention concerns formulation of one or more of the polynucleotide, polypeptide, T-cell and/or antibody compositions disclosed herein in pharmaceutically-acceptable solutions for administration to a cell or an animal, either alone, or in combination with one or more other modalities of therapy.

It will also be understood that, if desired, the nucleic acid segment, RNA, DNA or PNA compositions that express a polypeptide as disclosed herein may be administered in combination with other agents as well, such as, *e.g.*, other proteins or polypeptides or various pharmaceutically-active agents. In fact, there is virtually no limit to other components that may also be included, given that the additional agents do not cause a significant adverse effect upon contact with the target cells or host tissues. The compositions may thus be delivered along with various other agents as required in the particular instance. Such compositions may be purified from host cells or other biological sources, or alternatively may be chemically synthesized as described herein. Likewise,

such compositions may further comprise substituted or derivatized RNA or DNA compositions.

Formulation of pharmaceutically-acceptable excipients and carrier solutions is well-known to those of skill in the art, as is the development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens, including *e.g.*, oral, parenteral, intravenous, intranasal, and intramuscular administration and formulation.

1. ORAL DELIVERY

In certain applications, the pharmaceutical compositions disclosed herein may be delivered *via* oral administration to an animal. As such, these compositions may be formulated with an inert diluent or with an assimilable edible carrier, or they may be enclosed in hard- or soft-shell gelatin capsule, or they may be compressed into tablets, or they may be incorporated directly with the food of the diet.

The active compounds may even be incorporated with excipients and used in the form of ingestible tablets, buccal tables, troches, capsules, elixirs, suspensions, syrups, wafers, and the like (Mathiowitz *et al.*, 1997; Hwang *et al.*, 1998; U. S. Patent 5,641,515; U. S. Patent 5,580,579 and U. S. Patent 5,792,451, each specifically incorporated herein by reference in its entirety). The tablets, troches, pills, capsules and the like may also contain the following: a binder, as gum tragacanth, acacia, cornstarch, or gelatin; excipients, such as dicalcium phosphate; a disintegrating agent, such as corn starch, potato starch, alginic acid and the like; a lubricant, such as magnesium stearate; and a sweetening agent, such as sucrose, lactose or saccharin may be added or a flavoring agent, such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar, or both. A syrup of elixir may contain the active compound sucrose as a sweetening agent methyl and propylparabens as preservatives, a dye and flavoring, such as cherry or orange

flavor. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compounds may be incorporated into sustained-release preparation and formulations.

5 Typically, these formulations may contain at least about 0.1% of the active compound or more, although the percentage of the active ingredient(s) may, of course, be varied and may conveniently be between about 1 or 2% and about 60% or 70% or more of the weight or volume of the total formulation. Naturally, the amount of active compound(s) in each therapeutically useful composition may be prepared in such a way
10 that a suitable dosage will be obtained in any given unit dose of the compound. Factors such as solubility, bioavailability, biological half-life, route of administration, product shelf life, as well as other pharmacological considerations will be contemplated by one skilled in the art of preparing such pharmaceutical formulations, and as such, a variety of dosages and treatment regimens may be desirable.

15 For oral administration the compositions of the present invention may alternatively be incorporated with one or more excipients in the form of a mouthwash, dentifrice, buccal tablet, oral spray, or sublingual orally-administered formulation. For example, a mouthwash may be prepared incorporating the active ingredient in the required amount in an appropriate solvent, such as a sodium borate solution (Dobell's Solution).
20 Alternatively, the active ingredient may be incorporated into an oral solution such as one containing sodium borate, glycerin and potassium bicarbonate, or dispersed in a dentifrice, or added in a therapeutically-effective amount to a composition that may include water, binders, abrasives, flavoring agents, foaming agents, and humectants. Alternatively the compositions may be fashioned into a tablet or solution form that may be placed under the
25 tongue or otherwise dissolved in the mouth.

2. INJECTABLE DELIVERY

 In certain circumstances it will be desirable to deliver the pharmaceutical compositions disclosed herein parenterally, intravenously, intramuscularly, or even

intraperitoneally as described in U. S. Patent 5,543,158; U. S. Patent 5,641,515 and U. S. Patent 5,399,363 (each specifically incorporated herein by reference in its entirety). Solutions of the active compounds as free base or pharmacologically acceptable salts may be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose.

5 Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations contain a preservative to prevent the growth of microorganisms.

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions (U. S. Patent 5,466,468, specifically incorporated herein by reference in its entirety). In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing,

15 for example, water, ethanol, polyol (*e.g.*, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable oils. Proper fluidity may be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. The prevention of the action of microorganisms can be facilitated by various

20 antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

25 For parenteral administration in an aqueous solution, for example, the solution should be suitably buffered if necessary and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, a sterile aqueous medium that can be employed will be known to those of

skill in the art in light of the present disclosure. For example, one dosage may be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. The person responsible for administration will, in any event, determine the appropriate dose for the individual subject. Moreover, for human administration, preparations should meet sterility, pyrogenicity, and the general safety and purity standards as required by FDA Office of Biologics standards.

10 Sterile injectable solutions are prepared by incorporating the active compounds in the required amount in the appropriate solvent with various of the other ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized active ingredients into a sterile vehicle which contains the basic dispersion medium and the required other
15 ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum-drying and freeze-drying techniques which yield a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

The compositions disclosed herein may be formulated in a neutral or salt
20 form. Pharmaceutically-acceptable salts, include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or
25 ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like. Upon formulation, solutions will be administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective. The formulations are easily administered in a variety of dosage forms such as injectable solutions, drug-release capsules, and the like.

As used herein, "carrier" includes any and all solvents, dispersion media, vehicles, coatings, diluents, antibacterial and antifungal agents, isotonic and absorption delaying agents, buffers, carrier solutions, suspensions, colloids, and the like. The use of such media and agents for pharmaceutical active substances is well known in the art.

- 5 Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions.

The phrase "pharmaceutically-acceptable" refers to molecular entities and compositions that do not produce an allergic or similar untoward reaction when
 10 administered to a human. The preparation of an aqueous composition that contains a protein as an active ingredient is well understood in the art. Typically, such compositions are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid prior to injection can also be prepared. The preparation can also be emulsified.

15 3. NASAL DELIVERY

In certain embodiments, the pharmaceutical compositions may be delivered by intranasal sprays, inhalation, and/or other aerosol delivery vehicles. Methods for delivering genes, nucleic acids, and peptide compositions directly to the lungs *via* nasal aerosol sprays has been described *e.g.*, in U. S. Patent 5,756,353 and U. S. Patent 5,804,212
 20 (each specifically incorporated herein by reference in its entirety). Likewise, the delivery of drugs using intranasal microparticle resins (Takenaga *et al.*, 1998) and lysophosphatidyl-glycerol compounds (U. S. Patent 5,725,871, specifically incorporated herein by reference in its entirety) are also well-known in the pharmaceutical arts. Likewise, transmucosal drug delivery in the form of a polytetrafluoroethylene support matrix is described in U. S.
 25 Patent 5,780,045 (specifically incorporated herein by reference in its entirety).

4. LIPOSOME-, NANOCAPSULE-, AND MICROPARTICLE-MEDIATED DELIVERY

In certain embodiments, the inventors contemplate the use of liposomes, nanocapsules, microparticles, microspheres, lipid particles, vesicles, and the like, for the introduction of the compositions of the present invention into suitable host cells. In particular, the compositions of the present invention may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, or a nanoparticle or the like.

Such formulations may be preferred for the introduction of pharmaceutically-acceptable formulations of the nucleic acids or constructs disclosed herein. The formation and use of liposomes is generally known to those of skill in the art (see for example, Couvreur *et al.*, 1977; Couvreur, 1988; Lasic, 1998; which describes the use of liposomes and nanocapsules in the targeted antibiotic therapy for intracellular bacterial infections and diseases). Recently, liposomes were developed with improved serum stability and circulation half-times (Gabizon and Papahadjopoulos, 1988; Allen and Choun, 1987; U. S. Patent 5,741,516, specifically incorporated herein by reference in its entirety). Further, various methods of liposome and liposome like preparations as potential drug carriers have been reviewed (Takakura, 1998; Chandran *et al.*, 1997; Margalit, 1995; U. S. Patent 5,567,434; U. S. Patent 5,552,157; U. S. Patent 5,565,213; U. S. Patent 5,738,868 and U. S. Patent 5,795,587, each specifically incorporated herein by reference in its entirety).

Liposomes have been used successfully with a number of cell types that are normally resistant to transfection by other procedures including T cell suspensions, primary hepatocyte cultures and PC 12 cells (Renneisen *et al.*, 1990; Muller *et al.*, 1990). In addition, liposomes are free of the DNA length constraints that are typical of viral-based delivery systems. Liposomes have been used effectively to introduce genes, drugs (Heath and Martin, 1986; Heath *et al.*, 1986; Balazsovits *et al.*, 1989; Fresta and Puglisi, 1996), radiotherapeutic agents (Pikul *et al.*, 1987), enzymes (Imaizumi *et al.*, 1990a; Imaizumi *et al.*, 1990b), viruses (Faller and Baltimore, 1984), transcription factors and allosteric effectors (Nicolau and Gersonde, 1979) into a variety of cultured cell lines and animals. In

addition, several successful clinical trials examining the effectiveness of liposome-mediated drug delivery have been completed (Lopez-Berestein *et al.*, 1985a; 1985b; Coune, 1988; Sculier *et al.*, 1988). Furthermore, several studies suggest that the use of liposomes is not associated with autoimmune responses, toxicity or gonadal localization
 5 after systemic delivery (Mori and Fukatsu, 1992).

Liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs). MLVs generally have diameters of from 25 nm to 4 μ m. Sonication of MLVs results in the formation of small unilamellar vesicles (SUVs) with
 10 diameters in the range of 200 to 500 Å, containing an aqueous solution in the core.

Liposomes bear resemblance to cellular membranes and are contemplated for use in connection with the present invention as carriers for the peptide compositions. They are widely suitable as both water- and lipid-soluble substances can be entrapped, *i.e.* in the aqueous spaces and within the bilayer itself, respectively. It is possible that the drug-
 15 bearing liposomes may even be employed for site-specific delivery of active agents by selectively modifying the liposomal formulation.

In addition to the teachings of Couvreur *et al.* (1977; 1988), the following information may be utilized in generating liposomal formulations. Phospholipids can form a variety of structures other than liposomes when dispersed in water, depending on the
 20 molar ratio of lipid to water. At low ratios the liposome is the preferred structure. The physical characteristics of liposomes depend on pH, ionic strength and the presence of divalent cations. Liposomes can show low permeability to ionic and polar substances, but at elevated temperatures undergo a phase transition which markedly alters their permeability. The phase transition involves a change from a closely packed, ordered
 25 structure, known as the gel state, to a loosely packed, less-ordered structure, known as the fluid state. This occurs at a characteristic phase-transition temperature and results in an increase in permeability to ions, sugars and drugs.

In addition to temperature, exposure to proteins can alter the permeability of liposomes. Certain soluble proteins, such as cytochrome c, bind, deform and penetrate the

bilayer, thereby causing changes in permeability. Cholesterol inhibits this penetration of proteins, apparently by packing the phospholipids more tightly. It is contemplated that the most useful liposome formations for antibiotic and inhibitor delivery will contain cholesterol.

5 The ability to trap solutes varies between different types of liposomes. For example, MLVs are moderately efficient at trapping solutes, but SUVs are extremely inefficient. SUVs offer the advantage of homogeneity and reproducibility in size distribution, however, and a compromise between size and trapping efficiency is offered by large unilamellar vesicles (LUVs). These are prepared by ether evaporation and are three
10 to four times more efficient at solute entrapment than MLVs.

 In addition to liposome characteristics, an important determinant in entrapping compounds is the physicochemical properties of the compound itself. Polar compounds are trapped in the aqueous spaces and nonpolar compounds bind to the lipid bilayer of the vesicle. Polar compounds are released through permeation or when the
15 bilayer is broken, but nonpolar compounds remain affiliated with the bilayer unless it is disrupted by temperature or exposure to lipoproteins. Both types show maximum efflux rates at the phase transition temperature.

 Liposomes interact with cells *via* four different mechanisms: endocytosis by phagocytic cells of the reticuloendothelial system such as macrophages and neutrophils;
20 adsorption to the cell surface, either by nonspecific weak hydrophobic or electrostatic forces, or by specific interactions with cell-surface components; fusion with the plasma cell membrane by insertion of the lipid bilayer of the liposome into the plasma membrane, with simultaneous release of liposomal contents into the cytoplasm; and by transfer of liposomal
25 lipids to cellular or subcellular membranes, or vice versa, without any association of the liposome contents. It often is difficult to determine which mechanism is operative and more than one may operate at the same time.

 The fate and disposition of intravenously injected liposomes depend on their physical properties, such as size, fluidity, and surface charge. They may persist in tissues for h or days, depending on their composition, and half lives in the blood range from min to

several h. Larger liposomes, such as MLVs and LUVs, are taken up rapidly by phagocytic cells of the reticuloendothelial system, but physiology of the circulatory system restrains the exit of such large species at most sites. They can exit only in places where large openings or pores exist in the capillary endothelium, such as the sinusoids of the liver or spleen. Thus, these organs are the predominate site of uptake. On the other hand, SUVs show a broader tissue distribution but still are sequestered highly in the liver and spleen. In general, this *in vivo* behavior limits the potential targeting of liposomes to only those organs and tissues accessible to their large size. These include the blood, liver, spleen, bone marrow, and lymphoid organs.

Targeting is generally not a limitation in terms of the present invention. However, should specific targeting be desired, methods are available for this to be accomplished. Antibodies may be used to bind to the liposome surface and to direct the antibody and its drug contents to specific antigenic receptors located on a particular cell-type surface. Carbohydrate determinants (glycoprotein or glycolipid cell-surface components that play a role in cell-cell recognition, interaction and adhesion) may also be used as recognition sites as they have potential in directing liposomes to particular cell types. Mostly, it is contemplated that intravenous injection of liposomal preparations would be used, but other routes of administration are also conceivable.

Alternatively, the invention provides for pharmaceutically-acceptable nanocapsule formulations of the compositions of the present invention. Nanocapsules can generally entrap compounds in a stable and reproducible way (Henry-Michelland *et al.*, 1987; Quintanar-Guerrero *et al.*, 1998; Douglas *et al.*, 1987). To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1 μm) should be designed using polymers able to be degraded *in vivo*. Biodegradable polyalkyl-cyanoacrylate nanoparticles that meet these requirements are contemplated for use in the present invention. Such particles may be easily made, as described (Couvreux *et al.*, 1980; 1988; zur Muhlen *et al.*, 1998; Zambaux *et al.* 1998; Pinto-Alphandry *et al.*, 1995 and U. S. Patent 5,145,684, specifically incorporated herein by reference in its entirety).

IMMUNOGENIC COMPOSITIONS

In certain preferred embodiments of the present invention, immunogenic compositions, or vaccines, are provided. The immunogenic compositions will generally comprise one or more pharmaceutical compositions, such as those discussed above, in combination with an immunostimulant. An immunostimulant may be any substance that enhances or potentiates an immune response (antibody and/or cell-mediated) to an exogenous antigen. Examples of immunostimulants include adjuvants, biodegradable microspheres (*e.g.*, polylactic galactide) and liposomes (into which the compound is incorporated; *see e.g.*, Fullerton, U.S. Patent No. 4,235,877). Vaccine preparation is generally described in, for example, M.F. Powell and M.J. Newman, eds., "Vaccine Design (the subunit and adjuvant approach)," Plenum Press (NY, 1995). Pharmaceutical compositions and immunogenic compositions within the scope of the present invention may also contain other compounds, which may be biologically active or inactive. For example, one or more immunogenic portions of other tumor antigens may be present, either incorporated into a fusion polypeptide or as a separate compound, within the composition.

Illustrative immunogenic compositions may contain DNA encoding one or more of the polypeptides as described above, such that the polypeptide is generated *in situ*. As noted above, the DNA may be present within any of a variety of delivery systems known to those of ordinary skill in the art, including nucleic acid expression systems, bacteria and viral expression systems. Numerous gene delivery techniques are well known in the art, such as those described by Rolland, *Crit. Rev. Therap. Drug Carrier Systems* 15:143-198, 1998, and references cited therein. Appropriate nucleic acid expression systems contain the necessary DNA sequences for expression in the patient (such as a suitable promoter and terminating signal). Bacterial delivery systems involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an immunogenic portion of the polypeptide on its cell surface or secretes such an epitope. In a preferred embodiment, the DNA may be introduced using a viral expression system (*e.g.*, vaccinia or other pox virus, retrovirus, or adenovirus), which may involve the use of a non-pathogenic (defective), replication competent virus. Suitable systems are disclosed, for

example, in Fisher-Hoch *et al.*, *Proc. Natl. Acad. Sci. USA* 86:317-321, 1989; Flexner *et al.*, *Ann. N.Y. Acad. Sci.* 569:86-103, 1989; Flexner *et al.*, *Vaccine* 8:17-21, 1990; U.S. Patent Nos. 4,603,112, 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, *Biotechniques* 5 6:616-627, 1988; Rosenfeld *et al.*, *Science* 252:431-434, 1991; Kolls *et al.*, *Proc. Natl. Acad. Sci. USA* 91:215-219, 1994; Kass-Eisler *et al.*, *Proc. Natl. Acad. Sci. USA* 90:11498-11502, 1993; Guzman *et al.*, *Circulation* 88:2838-2848, 1993; and Guzman *et al.*, *Cir. Res.* 73:1202-1207, 1993. Techniques for incorporating DNA into such expression systems are well known to those of ordinary skill in the art. The DNA may also be 10 "naked," as described, for example, in Ulmer *et al.*, *Science* 259:1745-1749, 1993 and reviewed by Cohen, *Science* 259:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto biodegradable beads, which are efficiently transported into the cells. It will be apparent that an immunogenic composition may comprise both a polynucleotide and a polypeptide component. Such immunogenic compositions may 15 provide for an enhanced immune response.

It will be apparent that an immunogenic composition may contain pharmaceutically acceptable salts of the polynucleotides and polypeptides provided herein. Such salts may be prepared from pharmaceutically acceptable non-toxic bases, including organic bases (*e.g.*, salts of primary, secondary and tertiary amines and basic amino acids) 20 and inorganic bases (*e.g.*, sodium, potassium, lithium, ammonium, calcium and magnesium salts).

While any suitable carrier known to those of ordinary skill in the art may be employed in the compositions of this invention, the type of carrier will vary depending on the mode of administration. Compositions of the present invention may be formulated for 25 any appropriate manner of administration, including for example, topical, oral, nasal, intravenous, intracranial, intraperitoneal, subcutaneous or intramuscular administration. For parenteral administration, such as subcutaneous injection, the carrier preferably comprises water, saline, alcohol, a fat, a wax or a buffer. For oral administration, any of the above carriers or a solid carrier, such as mannitol, lactose, starch, magnesium stearate,

sodium saccharine, talcum, cellulose, glucose, sucrose, and magnesium carbonate, may be employed. Biodegradable microspheres (e.g., polylactate polyglycolate) may also be employed as carriers for the pharmaceutical compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S. Patent Nos. 4,897,268; 5,075,109; 5,928,647; 5,811,128; 5,820,883; 5,853,763; 5,814,344 and 5,942,252. One may also employ a carrier comprising the particulate-protein complexes described in U.S. Patent No. 5,928,647, which are capable of inducing a class I-restricted cytotoxic T lymphocyte responses in a host.

Such compositions may also comprise buffers (e.g., neutral buffered saline or phosphate buffered saline), carbohydrates (e.g., glucose, mannose, sucrose or dextrans), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants, bacteriostats, chelating agents such as EDTA or glutathione, adjuvants (e.g., aluminum hydroxide), solutes that render the formulation isotonic, hypotonic or weakly hypertonic with the blood of a recipient, suspending agents, thickening agents and/or preservatives. Alternatively, compositions of the present invention may be formulated as a lyophilizate. Compounds may also be encapsulated within liposomes using well known technology.

Any of a variety of immunostimulants may be employed in the immunogenic compositions of this invention. For example, an adjuvant may be included. Most adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bordetella pertussis* or *Mycobacterium tuberculosis* derived proteins. Suitable adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres; monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF or interleukin-2, -7, or -12, may also be used as adjuvants.

Within the immunogenic compositions provided herein, the adjuvant composition is preferably designed to induce an immune response predominantly of the Th1 type. High levels of Th1-type cytokines (*e.g.*, IFN- γ , TNF α , IL-2 and IL-12) tend to favor the induction of cell mediated immune responses to an administered antigen. In contrast, high levels of Th2-type cytokines (*e.g.*, IL-4, IL-5, IL-6 and IL-10) tend to favor the induction of humoral immune responses. Following application of an immunogenic composition as provided herein, a patient will support an immune response that includes Th1- and Th2-type responses. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-type cytokines will increase to a greater extent than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann and Coffman, *Ann. Rev. Immunol.* 7:145-173, 1989.

Preferred adjuvants for use in eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-de-O-acylated monophosphoryl lipid A (3D-MPL), together with an aluminum salt. MPL adjuvants are available from Corixa Corporation (Seattle, WA; *see* US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094). CpG-containing oligonucleotides (in which the CpG dinucleotide is unmethylated) also induce a predominantly Th1 response. Such oligonucleotides are well known and are described, for example, in WO 96/02555, WO 99/33488 and U.S. Patent Nos. 6,008,200 and 5,856,462. Immunostimulatory DNA sequences are also described, for example, by Sato *et al.*, *Science* 273:352, 1996. Another preferred adjuvant is a saponin, preferably QS21 (Aquila Biopharmaceuticals Inc., Framingham, MA), which may be used alone or in combination with other adjuvants. For example, an enhanced system involves the combination of a monophosphoryl lipid A and saponin derivative, such as the combination of QS21 and 3D-MPL as described in WO 94/00153, or a less reactogenic composition where the QS21 is quenched with cholesterol, as described in WO 96/33739. Other preferred formulations comprise an oil-in-water emulsion and tocopherol. A particularly potent adjuvant formulation involving QS21, 3D-MPL and tocopherol in an oil-in-water emulsion is described in WO 95/17210.

Other preferred adjuvants include Montanide ISA 720 (Seppic, France), SAF (Chiron, California, United States), ISCOMS (CSL), MF-59 (Chiron), the SBAS series of adjuvants (*e.g.*, SBAS-2 or SBAS-4, available from SmithKline Beecham, Rixensart, Belgium), Detox (Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) and
5 other aminoalkyl glucosaminide 4-phosphates (AGPs), such as those described in pending U.S. Patent Application Serial Nos. 08/853,826 and 09/074,720, the disclosures of which are incorporated herein by reference in their entireties.

Any immunogenic composition provided herein may be prepared using well known methods that result in a combination of antigen, immune response enhancer and a
10 suitable carrier or excipient. The compositions described herein may be administered as part of a sustained release formulation (*i.e.*, a formulation such as a capsule, sponge or gel (composed of polysaccharides, for example) that effects a slow release of compound following administration). Such formulations may generally be prepared using well known technology (*see, e.g.*, Coombes *et al.*, *Vaccine* 14:1429-1438, 1996) and administered by,
15 for example, oral, rectal or subcutaneous implantation, or by implantation at the desired target site. Sustained-release formulations may contain a polypeptide, polynucleotide or antibody dispersed in a carrier matrix and/or contained within a reservoir surrounded by a rate controlling membrane.

Carriers for use within such formulations are biocompatible, and may also
20 be biodegradable; preferably the formulation provides a relatively constant level of active component release. Such carriers include microparticles of poly(lactide-co-glycolide), polyacrylate, latex, starch, cellulose, dextran and the like. Other delayed-release carriers include supramolecular biovectors, which comprise a non-liquid hydrophilic core (*e.g.*, a cross-linked polysaccharide or oligosaccharide) and, optionally, an external layer
25 comprising an amphiphilic compound, such as a phospholipid (*see e.g.*, U.S. Patent No. 5,151,254 and PCT applications WO 94/20078, WO/94/23701 and WO 96/06638). The amount of active compound contained within a sustained release formulation depends upon the site of implantation, the rate and expected duration of release and the nature of the condition to be treated or prevented.

Any of a variety of delivery vehicles may be employed within pharmaceutical compositions and immunogenic compositions to facilitate production of an antigen-specific immune response that targets tumor cells. Delivery vehicles include antigen presenting cells (APCs), such as dendritic cells, macrophages, B cells, monocytes and other cells that may be engineered to be efficient APCs. Such cells may, but need not, be genetically modified to increase the capacity for presenting the antigen, to improve activation and/or maintenance of the T cell response, to have anti-tumor effects *per se* and/or to be immunologically compatible with the receiver (*i.e.*, matched HLA haplotype). APCs may generally be isolated from any of a variety of biological fluids and organs, including tumor and peritumoral tissues, and may be autologous, allogeneic, syngeneic or xenogeneic cells.

Certain preferred embodiments of the present invention use dendritic cells or progenitors thereof as antigen-presenting cells. Dendritic cells are highly potent APCs (Banchereau and Steinman, *Nature* 392:245-251, 1998) and have been shown to be effective as a physiological adjuvant for eliciting prophylactic or therapeutic antitumor immunity (*see* Timmerman and Levy, *Ann. Rev. Med.* 50:507-529, 1999). In general, dendritic cells may be identified based on their typical shape (stellate *in situ*, with marked cytoplasmic processes (dendrites) visible *in vitro*), their ability to take up, process and present antigens with high efficiency and their ability to activate naïve T cell responses. Dendritic cells may, of course, be engineered to express specific cell-surface receptors or ligands that are not commonly found on dendritic cells *in vivo* or *ex vivo*, and such modified dendritic cells are contemplated by the present invention. As an alternative to dendritic cells, secreted vesicles antigen-loaded dendritic cells (called exosomes) may be used within an immunogenic composition (*see* Zitvogel *et al.*, *Nature Med.* 4:594-600, 1998).

Dendritic cells and progenitors may be obtained from peripheral blood, bone marrow, tumor-infiltrating cells, peritumoral tissues-infiltrating cells, lymph nodes, spleen, skin, umbilical cord blood or any other suitable tissue or fluid. For example, dendritic cells may be differentiated *ex vivo* by adding a combination of cytokines such as GM-CSF, IL-4,

IL-13 and/or TNF α to cultures of monocytes harvested from peripheral blood. Alternatively, CD34 positive cells harvested from peripheral blood, umbilical cord blood or bone marrow may be differentiated into dendritic cells by adding to the culture medium combinations of GM-CSF, IL-3, TNF α , CD40 ligand, LPS, flt3 ligand and/or other
 5 compound(s) that induce differentiation, maturation and proliferation of dendritic cells.

Dendritic cells are conveniently categorized as "immature" and "mature" cells, which allows a simple way to discriminate between two well characterized phenotypes. However, this nomenclature should not be construed to exclude all possible intermediate stages of differentiation. Immature dendritic cells are characterized as APC
 10 with a high capacity for antigen uptake and processing, which correlates with the high expression of Fc γ receptor and mannose receptor. The mature phenotype is typically characterized by a lower expression of these markers, but a high expression of cell surface molecules responsible for T cell activation such as class I and class II MHC, adhesion molecules (*e.g.*, CD54 and CD11) and costimulatory molecules (*e.g.*, CD40, CD80, CD86
 15 and 4-1BB).

APCs may generally be transfected with a polynucleotide encoding a prostate-specific protein (or portion or other variant thereof) such that the prostate-specific polypeptide, or an immunogenic portion thereof, is expressed on the cell surface. Such transfection may take place *ex vivo*, and a composition comprising such transfected cells
 20 may then be used for therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be administered to a patient, resulting in transfection that occurs *in vivo*. *In vivo* and *ex vivo* transfection of dendritic cells, for example, may generally be performed using any methods known in the art, such as those described in WO 97/24447, or the gene gun approach
 25 described by Mahvi *et al.*, *Immunology and cell Biology* 75:456-460, 1997. Antigen loading of dendritic cells may be achieved by incubating dendritic cells or progenitor cells with the prostate-specific polypeptide, DNA (naked or within a plasmid vector) or RNA; or with antigen-expressing recombinant bacterium or viruses (*e.g.*, vaccinia, fowlpox, adenovirus or lentivirus vectors). Prior to loading, the polypeptide may be covalently

conjugated to an immunological partner that provides T cell help (*e.g.*, a carrier molecule). Alternatively, a dendritic cell may be pulsed with a non-conjugated immunological partner, separately or in the presence of the polypeptide.

Immunogenic compositions and pharmaceutical compositions may be presented in unit-dose or multi-dose containers, such as sealed ampoules or vials. Such containers are preferably hermetically sealed to preserve sterility of the formulation until use. In general, formulations may be stored as suspensions, solutions or emulsions in oily or aqueous vehicles. Alternatively, a immunogenic composition or pharmaceutical composition may be stored in a freeze-dried condition requiring only the addition of a sterile liquid carrier immediately prior to use.

CANCER THERAPY

In further aspects of the present invention, the compositions described herein may be used for immunotherapy of cancer, such as prostate cancer. Within such methods, pharmaceutical compositions and immunogenic compositions are typically administered to a patient. As used herein, a “patient” refers to any warm-blooded animal, preferably a human. A patient may or may not be afflicted with cancer. Accordingly, the above pharmaceutical compositions and immunogenic compositions may be used to prevent the development of a cancer or to treat a patient afflicted with a cancer. A cancer may be diagnosed using criteria generally accepted in the art, including the presence of a malignant tumor. Pharmaceutical compositions and immunogenic compositions may be administered either prior to or following surgical removal of primary tumors and/or treatment such as administration of radiotherapy or conventional chemotherapeutic drugs. Administration may be by any suitable method, including administration by intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal, intradermal, anal, vaginal, topical and oral routes.

Within certain embodiments, immunotherapy may be active immunotherapy, in which treatment relies on the *in vivo* stimulation of the endogenous host

immune system to react against tumors with the administration of immune response-modifying agents (such as polypeptides and polynucleotides as provided herein).

Within other embodiments, immunotherapy may be passive immunotherapy, in which treatment involves the delivery of agents with established tumor-immune reactivity (such as effector cells or antibodies) that can directly or indirectly mediate antitumor effects and does not necessarily depend on an intact host immune system. Examples of effector cells include T cells as discussed above, T lymphocytes (such as CD8⁺ cytotoxic T lymphocytes and CD4⁺ T-helper tumor-infiltrating lymphocytes), killer cells (such as Natural Killer cells and lymphokine-activated killer cells), B cells and antigen-presenting cells (such as dendritic cells and macrophages) expressing a polypeptide provided herein. T cell receptors and antibody receptors specific for the polypeptides recited herein may be cloned, expressed and transferred into other vectors or effector cells for adoptive immunotherapy. The polypeptides provided herein may also be used to generate antibodies or anti-idiotypic antibodies (as described above and in U.S. Patent No. 4,918,164) for passive immunotherapy.

Effector cells may generally be obtained in sufficient quantities for adoptive immunotherapy by growth *in vitro*, as described herein. Culture conditions for expanding single antigen-specific effector cells to several billion in number with retention of antigen recognition *in vivo* are well known in the art. Such *in vitro* culture conditions typically use intermittent stimulation with antigen, often in the presence of cytokines (such as IL-2) and non-dividing feeder cells. As noted above, immunoreactive polypeptides as provided herein may be used to rapidly expand antigen-specific T cell cultures in order to generate a sufficient number of cells for immunotherapy. In particular, antigen-presenting cells, such as dendritic, macrophage, monocyte, fibroblast and/or B cells, may be pulsed with immunoreactive polypeptides or transfected with one or more polynucleotides using standard techniques well known in the art. For example, antigen-presenting cells can be transfected with a polynucleotide having a promoter appropriate for increasing expression in a recombinant virus or other expression system. Cultured effector cells for use in therapy must be able to grow and distribute widely, and to survive long term *in vivo*.

Studies have shown that cultured effector cells can be induced to grow *in vivo* and to survive long term in substantial numbers by repeated stimulation with antigen supplemented with IL-2 (*see, for example, Cheever et al., Immunological Reviews 157:177, 1997*).

5 Alternatively, a vector expressing a polypeptide recited herein may be introduced into antigen presenting cells taken from a patient and clonally propagated *ex vivo* for transplant back into the same patient. Transfected cells may be reintroduced into the patient using any means known in the art, preferably in sterile form by intravenous, intracavitary, intraperitoneal or intratumor administration.

10 Routes and frequency of administration of the therapeutic compositions described herein, as well as dosage, will vary from individual to individual, and may be readily established using standard techniques. In general, the pharmaceutical compositions and immunogenic compositions may be administered by injection (*e.g.*, intracutaneous, intramuscular, intravenous or subcutaneous), intranasally (*e.g.*, by aspiration) or orally.

15 Preferably, between 1 and 10 doses may be administered over a 52 week period. Preferably, 6 doses are administered, at intervals of 1 month, and booster vaccinations may be given periodically thereafter. Alternate protocols may be appropriate for individual patients. A suitable dose is an amount of a compound that, when administered as described above, is capable of promoting an anti-tumor immune response, and is at least 10-50%

20 above the basal (*i.e.*, untreated) level. Such response can be monitored by measuring the anti-tumor antibodies in a patient or by vaccine-dependent generation of cytolytic effector cells capable of killing the patient's tumor cells *in vitro*. Such immunogenic compositions should also be capable of causing an immune response that leads to an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial or longer disease-free

25 survival) in treated patients as compared to non-treated patients. In general, for pharmaceutical compositions and immunogenic compositions comprising one or more polypeptides, the amount of each polypeptide present in a dose ranges from about 25 μ g to 5 mg per kg of host. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.

In general, an appropriate dosage and treatment regimen provides the active compound(s) in an amount sufficient to provide therapeutic and/or prophylactic benefit. Such a response can be monitored by establishing an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial, or longer disease-free survival) in treated patients as compared to non-treated patients. Increases in preexisting immune responses to a prostate-specific protein generally correlate with an improved clinical outcome. Such immune responses may generally be evaluated using standard proliferation, cytotoxicity or cytokine assays, which may be performed using samples obtained from a patient before and after treatment.

10 CANCER DETECTION AND DIAGNOSIS

In general, a cancer may be detected in a patient based on the presence of one or more prostate-specific proteins and/or polynucleotides encoding such proteins in a biological sample (for example, blood, sera, sputum urine and/or tumor biopsies) obtained from the patient. In other words, such proteins may be used as markers to indicate the presence or absence of a cancer such as prostate cancer. In addition, such proteins may be useful for the detection of other cancers. The binding agents provided herein generally permit detection of the level of antigen that binds to the agent in the biological sample. Polynucleotide primers and probes may be used to detect the level of mRNA encoding a tumor protein, which is also indicative of the presence or absence of a cancer. In general, a prostate-specific sequence should be present at a level that is at least three fold higher in prostate tissue than in other normal tissues.

There are a variety of assay formats known to those of ordinary skill in the art for using a binding agent to detect polypeptide markers in a sample. *See, e.g.*, Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, the presence or absence of a cancer in a patient may be determined by (a) contacting a biological sample obtained from a patient with a binding agent; (b) detecting in the sample a level of polypeptide that binds to the binding agent; and (c) comparing the level of polypeptide with a predetermined cut-off value.

In a preferred embodiment, the assay involves the use of binding agent immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a detection reagent that contains a reporter group and specifically binds to the binding agent/polypeptide complex.

5 Such detection reagents may comprise, for example, a binding agent that specifically binds to the polypeptide or an antibody or other agent that specifically binds to the binding agent, such as an anti-immunoglobulin, protein G, protein A or a lectin. Alternatively, a competitive assay may be utilized, in which a polypeptide is labeled with a reporter group and allowed to bind to the immobilized binding agent after incubation of the binding agent

10 with the sample. The extent to which components of the sample inhibit the binding of the labeled polypeptide to the binding agent is indicative of the reactivity of the sample with the immobilized binding agent. Suitable polypeptides for use within such assays include full length prostate-specific proteins and portions thereof to which the binding agent binds, as described above.

15 The solid support may be any material known to those of ordinary skill in the art to which the tumor protein may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a plastic material such as polystyrene or polyvinylchloride. The support may also be a magnetic

20 particle or a fiber optic sensor, such as those disclosed, for example, in U.S. Patent No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention, the term "immobilization" refers to both noncovalent association, such as adsorption, and covalent attachment (which

25 may be a direct linkage between the agent and functional groups on the support or may be a linkage by way of a cross-linking agent). Immobilization by adsorption to a well in a microtiter plate or to a membrane is preferred. In such cases, adsorption may be achieved by contacting the binding agent, in a suitable buffer, with the solid support for a suitable amount of time. The contact time varies with temperature, but is typically between about 1

hour and about 1 day. In general, contacting a well of a plastic microtiter plate (such as polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10 μ g, and preferably about 100 ng to about 1 μ g, is sufficient to immobilize an adequate amount of binding agent.

5 Covalent attachment of binding agent to a solid support may generally be achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports having an appropriate polymer coating using benzoquinone or by condensation of an aldehyde group
10 on the support with an amine and an active hydrogen on the binding partner (*see, e.g.*, Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized on a solid support, commonly the well of a microtiter plate, with the sample, such that
15 polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound sample is then removed from the immobilized polypeptide-antibody complexes and a detection reagent (preferably a second antibody capable of binding to a different site on the polypeptide) containing a reporter group is added. The amount of detection reagent that remains bound to the solid support is then determined using a method appropriate for the
20 specific reporter group.

More specifically, once the antibody is immobilized on the support as described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as bovine serum albumin or Tween 20™ (Sigma Chemical Co., St. Louis, MO). The immobilized
25 antibody is then incubated with the sample, and polypeptide is allowed to bind to the antibody. The sample may be diluted with a suitable diluent, such as phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact time (*i.e.*, incubation time) is a period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with prostate cancer. Preferably, the contact time is

sufficient to achieve a level of binding that is at least about 95% of that achieved at equilibrium between bound and unbound polypeptide. Those of ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about 30 minutes is generally sufficient.

Unbound sample may then be removed by washing the solid support with an appropriate buffer, such as PBS containing 0.1% Tween 20™. The second antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include those groups recited above.

The detection reagent is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide. An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound detection reagent is then removed and bound detection reagent is detected using the reporter group. The method employed for detecting the reporter group depends upon the nature of the reporter group. For radioactive groups, scintillation counting or autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group (commonly a radioactive or fluorescent group or an enzyme). Enzyme reporter groups may generally be detected by the addition of substrate (generally for a specific period of time), followed by spectroscopic or other analysis of the reaction products.

To determine the presence or absence of a cancer, such as prostate cancer, the signal detected from the reporter group that remains bound to the solid support is generally compared to a signal that corresponds to a predetermined cut-off value. In one preferred embodiment, the cut-off value for the detection of a cancer is the average mean signal obtained when the immobilized antibody is incubated with samples from patients without the cancer. In general, a sample generating a signal that is three standard deviations above the predetermined cut-off value is considered positive for the cancer. In an alternate preferred embodiment, the cut-off value is determined using a Receiver

Operator Curve, according to the method of Sackett *et al.*, *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, the cut-off value may be determined from a plot of pairs of true positive rates (*i.e.*, sensitivity) and false positive rates (100%-specificity) that correspond to each possible

5 cut-off value for the diagnostic test result. The cut-off value on the plot that is the closest to the upper left-hand corner (*i.e.*, the value that encloses the largest area) is the most accurate cut-off value, and a sample generating a signal that is higher than the cut-off value determined by this method may be considered positive. Alternatively, the cut-off value may be shifted to the left along the plot, to minimize the false positive rate, or to the right,

10 to minimize the false negative rate. In general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive for a cancer.

In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the binding agent is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the

15 immobilized binding agent as the sample passes through the membrane. A second, labeled binding agent then binds to the binding agent-polypeptide complex as a solution containing the second binding agent flows through the membrane. The detection of bound second binding agent may then be performed as described above. In the strip test format, one end of the membrane to which binding agent is bound is immersed in a solution containing the

20 sample. The sample migrates along the membrane through a region containing second binding agent and to the area of immobilized binding agent. Concentration of second binding agent at the area of immobilized antibody indicates the presence of a cancer. Typically, the concentration of second binding agent at that site generates a pattern, such as a line, that can be read visually. The absence of such a pattern indicates a negative result.

25 In general, the amount of binding agent immobilized on the membrane is selected to generate a visually discernible pattern when the biological sample contains a level of polypeptide that would be sufficient to generate a positive signal in the two-antibody sandwich assay, in the format discussed above. Preferred binding agents for use in such assays are antibodies and antigen-binding fragments thereof. Preferably, the amount of

antibody immobilized on the membrane ranges from about 25 ng to about 1 μ g, and more preferably from about 50 ng to about 500 ng. Such tests can typically be performed with a very small amount of biological sample.

Of course, numerous other assay protocols exist that are suitable for use with the tumor proteins or binding agents of the present invention. The above descriptions are intended to be exemplary only. For example, it will be apparent to those of ordinary skill in the art that the above protocols may be readily modified to use prostate-specific polypeptides to detect antibodies that bind to such polypeptides in a biological sample. The detection of such prostate-specific protein specific antibodies may correlate with the presence of a cancer.

A cancer may also, or alternatively, be detected based on the presence of T cells that specifically react with a prostate-specific protein in a biological sample. Within certain methods, a biological sample comprising CD4⁺ and/or CD8⁺ T cells isolated from a patient is incubated with a prostate-specific polypeptide, a polynucleotide encoding such a polypeptide and/or an APC that expresses at least an immunogenic portion of such a polypeptide, and the presence or absence of specific activation of the T cells is detected. Suitable biological samples include, but are not limited to, isolated T cells. For example, T cells may be isolated from a patient by routine techniques (such as by Ficoll/Hypaque density gradient centrifugation of peripheral blood lymphocytes). T cells may be incubated *in vitro* for 2-9 days (typically 4 days) at 37°C with polypeptide (*e.g.*, 5 - 25 μ g/ml). It may be desirable to incubate another aliquot of a T cell sample in the absence of prostate-specific polypeptide to serve as a control. For CD4⁺ T cells, activation is preferably detected by evaluating proliferation of the T cells. For CD8⁺ T cells, activation is preferably detected by evaluating cytolytic activity. A level of proliferation that is at least two fold greater and/or a level of cytolytic activity that is at least 20% greater than in disease-free patients indicates the presence of a cancer in the patient.

As noted above, a cancer may also, or alternatively, be detected based on the level of mRNA encoding a prostate-specific protein in a biological sample. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction

(PCR) based assay to amplify a portion of a prostate-specific cDNA derived from a biological sample, wherein at least one of the oligonucleotide primers is specific for (*i.e.*, hybridizes to) a polynucleotide encoding the prostate-specific protein. The amplified cDNA is then separated and detected using techniques well known in the art, such as gel electrophoresis. Similarly, oligonucleotide probes that specifically hybridize to a polynucleotide encoding a prostate-specific protein may be used in a hybridization assay to detect the presence of polynucleotide encoding the tumor protein in a biological sample.

To permit hybridization under assay conditions, oligonucleotide primers and probes should comprise an oligonucleotide sequence that has at least about 60%, preferably at least about 75% and more preferably at least about 90%, identity to a portion of a polynucleotide encoding a prostate-specific protein that is at least 10 nucleotides, and preferably at least 20 nucleotides, in length. Preferably, oligonucleotide primers and/or probes hybridize to a polynucleotide encoding a polypeptide described herein under moderately stringent conditions, as defined above. Oligonucleotide primers and/or probes which may be usefully employed in the diagnostic methods described herein preferably are at least 10-40 nucleotides in length. In a preferred embodiment, the oligonucleotide primers comprise at least 10 contiguous nucleotides, more preferably at least 15 contiguous nucleotides, of a DNA molecule having a sequence recited in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824. Techniques for both PCR based assays and hybridization assays are well known in the art (*see*, for example, Mullis *et al.*, *Cold Spring Harbor Symp. Quant. Biol.*, 51:263, 1987; Erlich ed., *PCR Technology*, Stockton Press, NY, 1989).

One preferred assay employs RT-PCR, in which PCR is applied in conjunction with reverse transcription. Typically, RNA is extracted from a biological sample, such as biopsy tissue, and is reverse transcribed to produce cDNA molecules. PCR amplification using at least one specific primer generates a cDNA molecule, which may be separated and visualized using, for example, gel electrophoresis. Amplification may be performed on biological samples taken from a test patient and from an individual who is

not afflicted with a cancer. The amplification reaction may be performed on several dilutions of cDNA spanning two orders of magnitude. A two-fold or greater increase in expression in several dilutions of the test patient sample as compared to the same dilutions of the non-cancerous sample is typically considered positive.

5 In another embodiment, the compositions described herein may be used as markers for the progression of cancer. In this embodiment, assays as described above for the diagnosis of a cancer may be performed over time, and the change in the level of reactive polypeptide(s) or polynucleotide(s) evaluated. For example, the assays may be performed every 24-72 hours for a period of 6 months to 1 year, and thereafter performed
10 as needed. In general, a cancer is progressing in those patients in whom the level of polypeptide or polynucleotide detected increases over time. In contrast, the cancer is not progressing when the level of reactive polypeptide or polynucleotide either remains constant or decreases with time.

Certain *in vivo* diagnostic assays may be performed directly on a tumor.
15 One such assay involves contacting tumor cells with a binding agent. The bound binding agent may then be detected directly or indirectly via a reporter group. Such binding agents may also be used in histological applications. Alternatively, polynucleotide probes may be used within such applications.

As noted above, to improve sensitivity, multiple prostate-specific protein
20 markers may be assayed within a given sample. It will be apparent that binding agents specific for different proteins provided herein may be combined within a single assay. Further, multiple primers or probes may be used concurrently. The selection of tumor protein markers may be based on routine experiments to determine combinations that results in optimal sensitivity. In addition, or alternatively, assays for tumor proteins
25 provided herein may be combined with assays for other known tumor antigens.

DIAGNOSTIC KITS

The present invention further provides kits for use within any of the above diagnostic methods. Such kits typically comprise two or more components necessary for

performing a diagnostic assay. Components may be compounds, reagents, containers and/or equipment. For example, one container within a kit may contain a monoclonal antibody or fragment thereof that specifically binds to a prostate-specific protein. Such antibodies or fragments may be provided attached to a support material, as described
 5 above. One or more additional containers may enclose elements, such as reagents or buffers, to be used in the assay. Such kits may also, or alternatively, contain a detection reagent as described above that contains a reporter group suitable for direct or indirect detection of antibody binding.

Alternatively, a kit may be designed to detect the level of mRNA encoding a
 10 prostate-specific protein in a biological sample. Such kits generally comprise at least one oligonucleotide probe or primer, as described above, that hybridizes to a polynucleotide encoding a prostate-specific protein. Such an oligonucleotide may be used, for example, within a PCR or hybridization assay. Additional components that may be present within such kits include a second oligonucleotide and/or a diagnostic reagent or container to
 15 facilitate the detection of a polynucleotide encoding a prostate-specific protein.

The following Examples are offered by way of illustration and not by way of limitation.

EXAMPLE 1

ISOLATION AND CHARACTERIZATION OF PROSTATE-SPECIFIC POLYPEPTIDES

This Example describes the isolation of certain prostate-specific
 5 polypeptides from a prostate tumor cDNA library.

A human prostate tumor cDNA expression library was constructed from prostate tumor poly A⁺ RNA using a Superscript Plasmid System for cDNA Synthesis and Plasmid Cloning kit (BRL Life Technologies, Gaithersburg, MD 20897) following the manufacturer's protocol. Specifically, prostate tumor tissues were homogenized with
 10 polytron (Kinematica, Switzerland) and total RNA was extracted using Trizol reagent (BRL Life Technologies) as directed by the manufacturer. The poly A⁺ RNA was then purified using a Qiagen oligotex spin column mRNA purification kit (Qiagen, Santa Clarita, CA 91355) according to the manufacturer's protocol. First-strand cDNA was synthesized using the NotI/Oligo-dT18 primer. Double-stranded cDNA was synthesized,
 15 ligated with EcoRI/BAXI adaptors (Invitrogen, San Diego, CA) and digested with NotI. Following size fractionation with Chroma Spin-1000 columns (Clontech, Palo Alto, CA), the cDNA was ligated into the EcoRI/NotI site of pCDNA3.1 (Invitrogen) and transformed into ElectroMax *E. coli* DH10B cells (BRL Life Technologies) by electroporation.

Using the same procedure, a normal human pancreas cDNA expression
 20 library was prepared from a pool of six tissue specimens (Clontech). The cDNA libraries were characterized by determining the number of independent colonies, the percentage of clones that carried insert, the average insert size and by sequence analysis. The prostate tumor library contained 1.64×10^7 independent colonies, with 70% of clones having an insert and the average insert size being 1745 base pairs. The normal pancreas cDNA
 25 library contained 3.3×10^6 independent colonies, with 69% of clones having inserts and the average insert size being 1120 base pairs. For both libraries, sequence analysis showed that the majority of clones had a full length cDNA sequence and were synthesized from mRNA, with minimal rRNA and mitochondrial DNA contamination.

cDNA library subtraction was performed using the above prostate tumor and normal pancreas cDNA libraries, as described by Hara *et al.* (*Blood*, 84:189-199, 1994) with some modifications. Specifically, a prostate tumor-specific subtracted cDNA library was generated as follows. Normal pancreas cDNA library (70 µg) was digested with
 5 EcoRI, NotI, and SfuI, followed by a filling-in reaction with DNA polymerase Klenow fragment. After phenol-chloroform extraction and ethanol precipitation, the DNA was dissolved in 100 µl of H₂O, heat-denatured and mixed with 100 µl (100 µg) of Photoprobe biotin (Vector Laboratories, Burlingame, CA). As recommended by the manufacturer, the resulting mixture was irradiated with a 270 W sunlamp on ice for 20 minutes. Additional
 10 Photoprobe biotin (50 µl) was added and the biotinylation reaction was repeated. After extraction with butanol five times, the DNA was ethanol-precipitated and dissolved in 23 µl H₂O to form the driver DNA.

To form the tracer DNA, 10 µg prostate tumor cDNA library was digested with BamHI and XhoI, phenol chloroform extracted and passed through Chroma spin-400
 15 columns (Clontech). Following ethanol precipitation, the tracer DNA was dissolved in 5 µl H₂O. Tracer DNA was mixed with 15 µl driver DNA and 20 µl of 2 x hybridization buffer (1.5 M NaCl/10 mM EDTA/50 mM HEPES pH 7.5/0.2% sodium dodecyl sulfate), overlaid with mineral oil, and heat-denatured completely. The sample was immediately transferred into a 68 °C water bath and incubated for 20 hours (long hybridization [LH]). The reaction
 20 mixture was then subjected to a streptavidin treatment followed by phenol/chloroform extraction. This process was repeated three more times. Subtracted DNA was precipitated, dissolved in 12 µl H₂O, mixed with 8 µl driver DNA and 20 µl of 2 x hybridization buffer, and subjected to a hybridization at 68 °C for 2 hours (short hybridization [SH]). After removal of biotinylated double-stranded DNA, subtracted cDNA was ligated into
 25 BamHI/XhoI site of chloramphenicol resistant pBCSK⁺ (Stratagene, La Jolla, CA 92037) and transformed into ElectroMax *E. coli* DH10B cells by electroporation to generate a prostate tumor specific subtracted cDNA library (referred to as “prostate subtraction 1”).

To analyze the subtracted cDNA library, plasmid DNA was prepared from 100 independent clones, randomly picked from the subtracted prostate tumor specific

library and grouped based on insert size. Representative cDNA clones were further characterized by DNA sequencing with a Perkin Elmer/Applied Biosystems Division Automated Sequencer Model 373A (Foster City, CA). Six cDNA clones, hereinafter referred to as F1-13, F1-12, F1-16, H1-1, H1-9 and H1-4, were shown to be abundant in the subtracted prostate-specific cDNA library. The determined 3' and 5' cDNA sequences for F1-12 are provided in SEQ ID NO: 2 and 3, respectively, with determined 3' cDNA sequences for F1-13, F1-16, H1-1, H1-9 and H1-4 being provided in SEQ ID NO: 1 and 4-7, respectively.

The cDNA sequences for the isolated clones were compared to known sequences in the gene bank using the EMBL and GenBank databases (release 96). Four of the prostate tumor cDNA clones, F1-13, F1-16, H1-1, and H1-4, were determined to encode the following previously identified proteins: prostate specific antigen (PSA), human glandular kallikrein, human tumor expression enhanced gene, and mitochondria cytochrome C oxidase subunit II. H1-9 was found to be identical to a previously identified human autonomously replicating sequence. No significant homologies to the cDNA sequence for F1-12 were found.

Subsequent studies led to the isolation of a full-length cDNA sequence for F1-12 (also referred to as P504S). This sequence is provided in SEQ ID NO: 107, with the corresponding predicted amino acid sequence being provided in SEQ ID NO: 108. cDNA splice variants of P504S are provided in SEQ ID NO: 600-605.

To clone less abundant prostate tumor specific genes, cDNA library subtraction was performed by subtracting the prostate tumor cDNA library described above with the normal pancreas cDNA library and with the three most abundant genes in the previously subtracted prostate tumor specific cDNA library: human glandular kallikrein, prostate specific antigen (PSA), and mitochondria cytochrome C oxidase subunit II. Specifically, 1 µg each of human glandular kallikrein, PSA and mitochondria cytochrome C oxidase subunit II cDNAs in pCDNA3.1 were added to the driver DNA and subtraction was performed as described above to provide a second subtracted cDNA library hereinafter referred to as the "subtracted prostate tumor specific cDNA library with spike".

Twenty-two cDNA clones were isolated from the subtracted prostate tumor specific cDNA library with spike. The determined 3' and 5' cDNA sequences for the clones referred to as J1-17, L1-12, N1-1862, J1-13, J1-19, J1-25, J1-24, K1-58, K1-63, L1-4 and L1-14 are provided in SEQ ID NOS: 8-9, 10-11, 12-13, 14-15, 16-17, 18-19, 20-21, 22-23, 24-25, 26-27 and 28-29, respectively. The determined 3' cDNA sequences for the clones referred to as J1-12, J1-16, J1-21, K1-48, K1-55, L1-2, L1-6, N1-1858, N1-1860, N1-1861, N1-1864 are provided in SEQ ID NOS: 30-40, respectively. Comparison of these sequences with those in the gene bank as described above, revealed no significant homologies to three of the five most abundant DNA species, (J1-17, L1-12 and N1-1862; SEQ ID NOS: 8-9, 10-11 and 12-13, respectively). Of the remaining two most abundant species, one (J1-12; SEQ ID NO:30) was found to be identical to the previously identified human pulmonary surfactant-associated protein, and the other (K1-48; SEQ ID NO:33) was determined to have some homology to *R. norvegicus* mRNA for 2-arylpropionyl-CoA epimerase. Of the 17 less abundant cDNA clones isolated from the subtracted prostate tumor specific cDNA library with spike, four (J1-16, K1-55, L1-6 and N1-1864; SEQ ID NOS:31, 34, 36 and 40, respectively) were found to be identical to previously identified sequences, two (J1-21 and N1-1860; SEQ ID NOS: 32 and 38, respectively) were found to show some homology to non-human sequences, and two (L1-2 and N1-1861; SEQ ID NOS: 35 and 39, respectively) were found to show some homology to known human sequences. No significant homologies were found to the polypeptides J1-13, J1-19, J1-24, J1-25, K1-58, K1-63, L1-4, L1-14 (SEQ ID NOS: 14-15, 16-17, 20-21, 18-19, 22-23, 24-25, 26-27, 28-29, respectively).

Subsequent studies led to the isolation of full length cDNA sequences for J1-17, L1-12 and N1-1862 (SEQ ID NOS: 109-111, respectively). The corresponding predicted amino acid sequences are provided in SEQ ID NOS: 112-114. L1-12 is also referred to as P501S. A cDNA splice variant of P501S is provided in SEQ ID NO: 606.

In a further experiment, four additional clones were identified by subtracting a prostate tumor cDNA library with normal prostate cDNA prepared from a pool of three normal prostate poly A⁺ RNA (referred to as "prostate subtraction 2"). The determined

cDNA sequences for these clones, hereinafter referred to as U1-3064, U1-3065, V1-3692 and 1A-3905, are provided in SEQ ID NO: 69-72, respectively. Comparison of the determined sequences with those in the gene bank revealed no significant homologies to U1-3065.

5 A second subtraction with spike (referred to as “prostate subtraction spike 2”) was performed by subtracting a prostate tumor specific cDNA library with spike with normal pancreas cDNA library and further spiked with PSA, J1-17, pulmonary surfactant-associated protein, mitochondrial DNA, cytochrome c oxidase subunit II, N1-1862, autonomously replicating sequence, L1-12 and tumor expression enhanced gene. Four
10 additional clones, hereinafter referred to as V1-3686, R1-2330, 1B-3976 and V1-3679, were isolated. The determined cDNA sequences for these clones are provided in SEQ ID NO:73-76, respectively. Comparison of these sequences with those in the gene bank revealed no significant homologies to V1-3686 and R1-2330.

 Further analysis of the three prostate subtractions described above (prostate
15 subtraction 2, subtracted prostate tumor specific cDNA library with spike, and prostate subtraction spike 2) resulted in the identification of sixteen additional clones, referred to as 1G-4736, 1G-4738, 1G-4741, 1G-4744, 1G-4734, 1H-4774, 1H-4781, 1H-4785, 1H-4787, 1H-4796, 1I-4810, 1I-4811, 1J-4876, 1K-4884 and 1K-4896. The determined cDNA sequences for these clones are provided in SEQ ID NOS: 77-92, respectively. Comparison
20 of these sequences with those in the gene bank as described above, revealed no significant homologies to 1G-4741, 1G-4734, 1I-4807, 1J-4876 and 1K-4896 (SEQ ID NOS: 79, 81, 87, 90 and 92, respectively). Further analysis of the isolated clones led to the determination of extended cDNA sequences for 1G-4736, 1G-4738, 1G-4741, 1G-4744, 1H-4774, 1H-4781, 1H-4785, 1H-4787, 1H-4796, 1I-4807, 1J-4876, 1K-4884 and 1K-
25 4896, provided in SEQ ID NOS: 179-188 and 191-193, respectively, and to the determination of additional partial cDNA sequences for 1I-4810 and 1I-4811, provided in SEQ ID NOS: 189 and 190, respectively.

 Additional studies with prostate subtraction spike 2 resulted in the isolation of three more clones. Their sequences were determined as described above and compared

to the most recent GenBank. All three clones were found to have homology to known genes, which are Cysteine-rich protein, KIAA0242, and KIAA0280 (SEQ ID NO: 317, 319, and 320, respectively). Further analysis of these clones by Synteni microarray (Synteni, Palo Alto, CA) demonstrated that all three clones were over-expressed in most prostate tumors and prostate BPH, as well as in the majority of normal prostate tissues tested, but low expression in all other normal tissues.

An additional subtraction was performed by subtracting a normal prostate cDNA library with normal pancreas cDNA (referred to as "prostate subtraction 3"). This led to the identification of six additional clones referred to as 1G-4761, 1G-4762, 1H-4766, 1H-4770, 1H-4771 and 1H-4772 (SEQ ID NOS: 93-98). Comparison of these sequences with those in the gene bank revealed no significant homologies to 1G-4761 and 1H-4771 (SEQ ID NOS: 93 and 97, respectively). Further analysis of the isolated clones led to the determination of extended cDNA sequences for 1G-4761, 1G-4762, 1H-4766 and 1H-4772 provided in SEQ ID NOS: 194-196 and 199, respectively, and to the determination of additional partial cDNA sequences for 1H-4770 and 1H-4771, provided in SEQ ID NOS: 197 and 198, respectively.

Subtraction of a prostate tumor cDNA library, prepared from a pool of polyA⁺ RNA from three prostate cancer patients, with a normal pancreas cDNA library (prostate subtraction 4) led to the identification of eight clones, referred to as 1D-4297, 1D-4309, 1D-4278, 1D-4288, 1D-4283, 1D-4304, 1D-4296 and 1D-4280 (SEQ ID NOS: 99-107). These sequences were compared to those in the gene bank as described above. No significant homologies were found to 1D-4283 and 1D-4304 (SEQ ID NOS: 103 and 104, respectively). Further analysis of the isolated clones led to the determination of extended cDNA sequences for 1D-4309, 1D-4278, 1D-4288, 1D-4283, 1D-4304, 1D-4296 and 1D-4280, provided in SEQ ID NOS: 200-206, respectively.

cDNA clones isolated in prostate subtraction 1 and prostate subtraction 2, described above, were colony PCR amplified and their mRNA expression levels in prostate tumor, normal prostate and in various other normal tissues were determined using microarray technology (Synteni, Palo Alto, CA). Briefly, the PCR amplification products

were dotted onto slides in an array format, with each product occupying a unique location in the array. mRNA was extracted from the tissue sample to be tested, reverse transcribed, and fluorescent-labeled cDNA probes were generated. The microarrays were probed with the labeled cDNA probes, the slides scanned and fluorescence intensity was measured.

- 5 This intensity correlates with the hybridization intensity. Two clones (referred to as P509S and P510S) were found to be over-expressed in prostate tumor and normal prostate and expressed at low levels in all other normal tissues tested (liver, pancreas, skin, bone marrow, brain, breast, adrenal gland, bladder, testes, salivary gland, large intestine, kidney, ovary, lung, spinal cord, skeletal muscle and colon). The determined cDNA sequences for
- 10 P509S and P510S are provided in SEQ ID NO: 223 and 224, respectively. Comparison of these sequences with those in the gene bank as described above, revealed some homology to previously identified ESTs.

- Additional, studies led to the isolation of the full-length cDNA sequence for P509S. This sequence is provided in SEQ ID NO: 332, with the corresponding predicted
- 15 amino acid sequence being provided in SEQ ID NO: 339. Two variant full-length cDNA sequences for P510S are provided in SEQ ID NO: 535 and 536, with the corresponding predicted amino acid sequences being provided in SEQ ID NO: 537 and 538, respectively. Additional splice variants of P510S are provided in SEQ ID NO: 598 and 599.

- The determined cDNA sequences for additional prostate-specific clones
- 20 isolated during characterization of prostate specific cDNA libraries are provided in SEQ ID NO: 618-689, 691-697 and 709-772. Comparison of these sequences with those in the public databases revealed no significant homologies to any of these sequences.

EXAMPLE 2

- 25 DETERMINATION OF TISSUE SPECIFICITY OF PROSTATE-SPECIFIC POLYPEPTIDES

Using gene specific primers, mRNA expression levels for the representative prostate-specific polypeptides F1-16, H1-1, J1-17 (also referred to as P502S), L1-12 (also

referred to as P501S), F1-12 (also referred to as P504S) and N1-1862 (also referred to as P503S) were examined in a variety of normal and tumor tissues using RT-PCR.

Briefly, total RNA was extracted from a variety of normal and tumor tissues using Trizol reagent as described above. First strand synthesis was carried out using 1-2
 5 μ g of total RNA with SuperScript II reverse transcriptase (BRL Life Technologies) at 42 °C for one hour. The cDNA was then amplified by PCR with gene-specific primers. To ensure the semi-quantitative nature of the RT-PCR, β -actin was used as an internal control for each of the tissues examined. First, serial dilutions of the first strand cDNAs were prepared and RT-PCR assays were performed using β -actin specific primers. A dilution
 10 was then chosen that enabled the linear range amplification of the β -actin template and which was sensitive enough to reflect the differences in the initial copy numbers. Using these conditions, the β -actin levels were determined for each reverse transcription reaction from each tissue. DNA contamination was minimized by DNase treatment and by assuring a negative PCR result when using first strand cDNA that was prepared without adding
 15 reverse transcriptase.

mRNA Expression levels were examined in four different types of tumor tissue (prostate tumor from 2 patients, breast tumor from 3 patients, colon tumor, lung tumor), and sixteen different normal tissues, including prostate, colon, kidney, liver, lung, ovary, pancreas, skeletal muscle, skin, stomach, testes, bone marrow and brain. F1-16 was
 20 found to be expressed at high levels in prostate tumor tissue, colon tumor and normal prostate, and at lower levels in normal liver, skin and testes, with expression being undetectable in the other tissues examined. H1-1 was found to be expressed at high levels in prostate tumor, lung tumor, breast tumor, normal prostate, normal colon and normal brain, at much lower levels in normal lung, pancreas, skeletal muscle, skin, small intestine,
 25 bone marrow, and was not detected in the other tissues tested. J1-17 (P502S) and L1-12 (P501S) appear to be specifically over-expressed in prostate, with both genes being expressed at high levels in prostate tumor and normal prostate but at low to undetectable levels in all the other tissues examined. N1-1862 (P503S) was found to be over-expressed in 60% of prostate tumors and detectable in normal colon and kidney. The RT-PCR results

thus indicate that F1-16, H1-1, J1-17 (P502S), N1-1862 (P503S) and L1-12 (P501S) are either prostate specific or are expressed at significantly elevated levels in prostate.

Further RT-PCR studies showed that F1-12 (P504S) is over-expressed in 60% of prostate tumors, detectable in normal kidney but not detectable in all other tissues tested. Similarly, R1-2330 was shown to be over-expressed in 40% of prostate tumors, detectable in normal kidney and liver, but not detectable in all other tissues tested. U1-3064 was found to be over-expressed in 60% of prostate tumors, and also expressed in breast and colon tumors, but was not detectable in normal tissues.

RT-PCR characterization of R1-2330, U1-3064 and 1D-4279 showed that these three antigens are over-expressed in prostate and/or prostate tumors.

Northern analysis with four prostate tumors, two normal prostate samples, two BPH prostates, and normal colon, kidney, liver, lung, pancreas, skeletal muscle, brain, stomach, testes, small intestine and bone marrow, showed that L1-12 (P501S) is over-expressed in prostate tumors and normal prostate, while being undetectable in other normal tissues tested. J1-17 (P502S) was detected in two prostate tumors and not in the other tissues tested. N1-1862 (P503S) was found to be over-expressed in three prostate tumors and to be expressed in normal prostate, colon and kidney, but not in other tissues tested. F1-12 (P504S) was found to be highly expressed in two prostate tumors and to be undetectable in all other tissues tested.

The microarray technology described above was used to determine the expression levels of representative antigens described herein in prostate tumor, breast tumor and the following normal tissues: prostate, liver, pancreas, skin, bone marrow, brain, breast, adrenal gland, bladder, testes, salivary gland, large intestine, kidney, ovary, lung, spinal cord, skeletal muscle and colon. L1-12 (P501S) was found to be over-expressed in normal prostate and prostate tumor, with some expression being detected in normal skeletal muscle. Both J1-12 and F1-12 (P504S) were found to be over-expressed in prostate tumor, with expression being lower or undetectable in all other tissues tested. N1-1862 (P503S) was found to be expressed at high levels in prostate tumor and normal prostate, and at low levels in normal large intestine and normal colon, with expression

being undetectable in all other tissues tested. R1-2330 was found to be over-expressed in prostate tumor and normal prostate, and to be expressed at lower levels in all other tissues tested. 1D-4279 was found to be over-expressed in prostate tumor and normal prostate, expressed at lower levels in normal spinal cord, and to be undetectable in all other tissues
5 tested.

Further microarray analysis to specifically address the extent to which P501S (SEQ ID NO: 110) was expressed in breast tumor revealed moderate over-expression not only in breast tumor, but also in metastatic breast tumor (2/31), with negligible to low expression in normal tissues. This data suggests that P501S may be over-
10 expressed in various breast tumors as well as in prostate tumors.

The expression levels of 32 ESTs (expressed sequence tags) described by Vasmatazis *et al.* (*Proc. Natl. Acad. Sci. USA* 95:300-304, 1998) in a variety of tumor and normal tissues were examined by microarray technology as described above. Two of these clones (referred to as P1000C and P1001C) were found to be over-expressed in prostate
15 tumor and normal prostate, and expressed at low to undetectable levels in all other tissues tested (normal aorta, thymus, resting and activated PBMC, epithelial cells, spinal cord, adrenal gland, fetal tissues, skin, salivary gland, large intestine, bone marrow, liver, lung, dendritic cells, stomach, lymph nodes, brain, heart, small intestine, skeletal muscle, colon and kidney. The determined cDNA sequences for P1000C and P1001C are provided in
20 SEQ ID NO: 384 and 472, respectively. The sequence of P1001C was found to show some homology to the previously isolated Human mRNA for JM27 protein. No significant homologies were found to the sequence of P1000C.

The expression of the polypeptide encoded by the full length cDNA sequence for F1-12 (also referred to as P504S; SEQ ID NO: 108) was investigated by
25 immunohistochemical analysis. Rabbit-anti-P504S polyclonal antibodies were generated against the full length P504S protein by standard techniques. Subsequent isolation and characterization of the polyclonal antibodies were also performed by techniques well known in the art. Immunohistochemical analysis showed that the P504S polypeptide was expressed in 100% of prostate carcinoma samples tested (n=5).

The rabbit-anti-P504S polyclonal antibody did not appear to label benign prostate cells with the same cytoplasmic granular staining, but rather with light nuclear staining. Analysis of normal tissues revealed that the encoded polypeptide was found to be expressed in some, but not all normal human tissues. Positive cytoplasmic staining with
 5 rabbit-anti-P504S polyclonal antibody was found in normal human kidney, liver, brain, colon and lung-associated macrophages, whereas heart and bone marrow were negative.

This data indicates that the P504S polypeptide is present in prostate cancer tissues, and that there are qualitative and quantitative differences in the staining between benign prostatic hyperplasia tissues and prostate cancer tissues, suggesting that this
 10 polypeptide may be detected selectively in prostate tumors and therefore be useful in the diagnosis of prostate cancer.

EXAMPLE 3

ISOLATION AND CHARACTERIZATION OF PROSTATE-SPECIFIC 15 POLYPEPTIDES BY PCR-BASED SUBTRACTION

A cDNA subtraction library, containing cDNA from normal prostate subtracted with ten other normal tissue cDNAs (brain, heart, kidney, liver, lung, ovary, placenta, skeletal muscle, spleen and thymus) and then submitted to a first round of PCR
 20 amplification, was purchased from Clontech. This library was subjected to a second round of PCR amplification, following the manufacturer's protocol. The resulting cDNA fragments were subcloned into the vector pT7 Blue T-vector (Novagen, Madison, WI) and transformed into XL-1 Blue MRF' *E. coli* (Stratagene). DNA was isolated from independent clones and sequenced using a Perkin Elmer/Applied Biosystems Division
 25 Automated Sequencer Model 373A.

Fifty-nine positive clones were sequenced. Comparison of the DNA sequences of these clones with those in the gene bank, as described above, revealed no significant homologies to 25 of these clones, hereinafter referred to as P5, P8, P9, P18, P20, P30, P34, P36, P38, P39, P42, P49, P50, P53, P55, P60, P64, P65, P73, P75, P76, P79

and P84. The determined cDNA sequences for these clones are provided in SEQ ID NO: 41-45, 47-52 and 54-65, respectively. P29, P47, P68, P80 and P82 (SEQ ID NO: 46, 53 and 66-68, respectively) were found to show some degree of homology to previously identified DNA sequences. To the best of the inventors' knowledge, none of these sequences have been previously shown to be present in prostate.

Further studies employing the sequence of SEQ ID NO: 67 as a probe in standard full-length cloning methods, resulted in the isolation of three cDNA sequences which appear to be splice variants of P80 (also known as P704P). These sequences are provided in SEQ ID NO: 699-701.

Further studies using the PCR-based methodology described above resulted in the isolation of more than 180 additional clones, of which 23 clones were found to show no significant homologies to known sequences. The determined cDNA sequences for these clones are provided in SEQ ID NO: 115-123, 127, 131, 137, 145, 147-151, 153, 156-158 and 160. Twenty-three clones (SEQ ID NO: 124-126, 128-130, 132-136, 138-144, 146, 152, 154, 155 and 159) were found to show some homology to previously identified ESTs. An additional ten clones (SEQ ID NO: 161-170) were found to have some degree of homology to known genes. Larger cDNA clones containing the P20 sequence represent splice variants of a gene referred to as P703P. The determined DNA sequence for the variants referred to as DE1, DE13 and DE14 are provided in SEQ ID NOS: 171, 175 and 177, respectively, with the corresponding predicted amino acid sequences being provided in SEQ ID NO: 172, 176 and 178, respectively. The determined cDNA sequence for an extended spliced form of P703 is provided in SEQ ID NO: 225. The DNA sequences for the splice variants referred to as DE2 and DE6 are provided in SEQ ID NOS: 173 and 174, respectively.

mRNA Expression levels for representative clones in tumor tissues (prostate (n=5), breast (n=2), colon and lung) normal tissues (prostate (n=5), colon, kidney, liver, lung (n=2), ovary (n=2), skeletal muscle, skin, stomach, small intestine and brain), and activated and non-activated PBMC was determined by RT-PCR as described above. Expression was examined in one sample of each tissue type unless otherwise indicated.

P9 was found to be highly expressed in normal prostate and prostate tumor compared to all normal tissues tested except for normal colon which showed comparable expression. P20, a portion of the P703P gene, was found to be highly expressed in normal prostate and prostate tumor, compared to all twelve normal tissues tested. A modest increase in expression of P20 in breast tumor (n=2), colon tumor and lung tumor was seen compared to all normal tissues except lung (1 of 2). Increased expression of P18 was found in normal prostate, prostate tumor and breast tumor compared to other normal tissues except lung and stomach. A modest increase in expression of P5 was observed in normal prostate compared to most other normal tissues. However, some elevated expression was seen in normal lung and PBMC. Elevated expression of P5 was also observed in prostate tumors (2 of 5), breast tumor and one lung tumor sample. For P30, similar expression levels were seen in normal prostate and prostate tumor, compared to six of twelve other normal tissues tested. Increased expression was seen in breast tumors, one lung tumor sample and one colon tumor sample, and also in normal PBMC. P29 was found to be over-expressed in prostate tumor (5 of 5) and normal prostate (5 of 5) compared to the majority of normal tissues. However, substantial expression of P29 was observed in normal colon and normal lung (2 of 2). P80 was found to be over-expressed in prostate tumor (5 of 5) and normal prostate (5 of 5) compared to all other normal tissues tested, with increased expression also being seen in colon tumor.

Further studies resulted in the isolation of twelve additional clones, hereinafter referred to as 10-d8, 10-h10, 11-c8, 7-g6, 8-b5, 8-b6, 8-d4, 8-d9, 8-g3, 8-h11, 9-f12 and 9-f3. The determined DNA sequences for 10-d8, 10-h10, 11-c8, 8-d4, 8-d9, 8-h11, 9-f12 and 9-f3 are provided in SEQ ID NO: 207, 208, 209, 216, 217, 220, 221 and 222, respectively. The determined forward and reverse DNA sequences for 7-g6, 8-b5, 8-b6 and 8-g3 are provided in SEQ ID NO: 210 and 211; 212 and 213; 214 and 215; and 218 and 219, respectively. Comparison of these sequences with those in the gene bank revealed no significant homologies to the sequence of 9-f3. The clones 10-d8, 11-c8 and 8-h11 were found to show some homology to previously isolated ESTs, while 10-h10, 8-b5, 8-b6, 8-d4, 8-d9, 8-g3 and 9-f12 were found to show some homology to previously identified genes.

Further characterization of 7-G6 and 8-G3 showed identity to the known genes PAP and PSA, respectively.

mRNA expression levels for these clones were determined using the micro-array technology described above. The clones 7-G6, 8-G3, 8-B5, 8-B6, 8-D4, 8-D9, 9-F3, 9-F12, 9-H3, 10-A2, 10-A4, 11-C9 and 11-F2 were found to be over-expressed in prostate tumor and normal prostate, with expression in other tissues tested being low or undetectable. Increased expression of 8-F11 was seen in prostate tumor and normal prostate, bladder, skeletal muscle and colon. Increased expression of 10-H10 was seen in prostate tumor and normal prostate, bladder, lung, colon, brain and large intestine. Increased expression of 9-B1 was seen in prostate tumor, breast tumor, and normal prostate, salivary gland, large intestine and skin, with increased expression of 11-C8 being seen in prostate tumor, and normal prostate and large intestine.

An additional cDNA fragment derived from the PCR-based normal prostate subtraction, described above, was found to be prostate specific by both micro-array technology and RT-PCR. The determined cDNA sequence of this clone (referred to as 9-A11) is provided in SEQ ID NO: 226. Comparison of this sequence with those in the public databases revealed 99% identity to the known gene HOXB13.

Further studies led to the isolation of the clones 8-C6 and 8-H7. The determined cDNA sequences for these clones are provided in SEQ ID NO: 227 and 228, respectively. These sequences were found to show some homology to previously isolated ESTs.

PCR and hybridization-based methodologies were employed to obtain longer cDNA sequences for clone P20 (also referred to as P703P), yielding three additional cDNA fragments that progressively extend the 5' end of the gene. These fragments, referred to as P703PDE5, P703P6.26, and P703PX-23 (SEQ ID NO: 326, 328 and 330, with the predicted corresponding amino acid sequences being provided in SEQ ID NO: 327, 329 and 331, respectively) contain additional 5' sequence. P703PDE5 was recovered by screening of a cDNA library (#141-26) with a portion of P703P as a probe. P703P6.26 was recovered from a mixture of three prostate tumor cDNAs and P703PX_23 was

recovered from cDNA library (#438-48). Together, the additional sequences include all of the putative mature serine protease along with part of the putative signal sequence. The full-length cDNA sequence for P703P is provided in SEQ ID NO: 524, with the corresponding amino acid sequence being provided in SEQ ID NO: 525.

5 P703P was found to show some homology to previously identified proteases, such as thrombin. The thrombin receptor has been shown to be preferentially expressed in highly metastatic breast carcinoma cells and breast carcinoma biopsy samples. Introduction of thrombin receptor antisense cDNA has been shown to inhibit the invasion of metastatic breast carcinoma cells in culture. Antibodies against thrombin receptor
10 inhibit thrombin receptor activation and thrombin-induced platelet activation. Furthermore, peptides that resemble the receptor's tethered ligand domain inhibit platelet aggregation by thrombin. P703P may play a role in prostate cancer through a protease-activated receptor on the cancer cell or on stromal cells. The potential trypsin-like protease activity of P703P may either activate a protease-activated receptor on the cancer cell
15 membrane to promote tumorigenesis or activate a protease-activated receptor on the adjacent cells (such as stromal cells) to secrete growth factors and/or proteases (such as matrix metalloproteinases) that could promote tumor angiogenesis, invasion and metastasis. P703P may thus promote tumor progression and/or metastasis through the activation of protease-activated receptor. Polypeptides and antibodies that block the
20 P703P-receptor interaction may therefore be usefully employed in the treatment of prostate cancer.

Further studies using a PCR-based subtraction library of a prostate tumor pool subtracted against a pool of normal tissues (referred to as JP: PCR subtraction) resulted in the isolation of thirteen additional clones, seven of which did not share any
25 significant homology to known GenBank sequences. The determined cDNA sequences for these seven clones (P711P, P712P, novel 23, P774P, P775P, P710P and P768P) are provided in SEQ ID NO: 307-311, 313 and 315, respectively. The remaining six clones (SEQ ID NO: 316 and 321-325) were shown to share some homology to known genes. By microarray analysis, all thirteen clones showed three or more fold over-expression in

prostate tissues, including prostate tumors, BPH and normal prostate as compared to normal non-prostate tissues. Clones P711P, P712P, novel 23 and P768P showed over-expression in most prostate tumors and BPH tissues tested (n=29), and in the majority of normal prostate tissues (n=4), but background to low expression levels in all normal
 5 tissues. Clones P774P, P775P and P710P showed comparatively lower expression and expression in fewer prostate tumors and BPH samples, with negative to low expression in normal prostate.

Further studies led to the isolation of an extended cDNA sequence for P712P (SEQ ID NO: 552). The amino acid sequences encoded by 16 predicted open
 10 reading frames present within the sequence of SEQ ID NO: 552 are provided in SEQ ID NO: 553-568.

The full-length cDNA for P711P was obtained by employing the partial sequence of SEQ ID NO: 307 to screen a prostate cDNA library. Specifically, a directionally cloned prostate cDNA library was prepared using standard techniques. One
 15 million colonies of this library were plated onto LB/Amp plates. Nylon membrane filters were used to lift these colonies, and the cDNAs which were picked up by these filters were denatured and cross-linked to the filters by UV light. The P711P cDNA fragment of SEQ ID NO: 307 was radio-labeled and used to hybridize with these filters. Positive clones were selected, and cDNAs were prepared and sequenced using an automatic Perkin
 20 Elmer/Applied Biosystems sequencer. The determined full-length sequence of P711P is provided in SEQ ID NO: 382, with the corresponding predicted amino acid sequence being provided in SEQ ID NO: 383.

Using PCR and hybridization-based methodologies, additional cDNA sequence information was derived for two clones described above, 11-C9 and 9-F3, herein
 25 after referred to as P707P and P714P, respectively (SEQ ID NO: 333 and 334). After comparison with the most recent GenBank, P707P was found to be a splice variant of the known gene HoxB13. In contrast, no significant homologies to P714P were found. Further studies employing the sequence of SEQ ID NO: 334 as a probe in standard full-length cloning methods, resulted in an extended cDNA sequence for P714P. This sequence

is provided in SEQ ID NO: 698. This sequence was found to show some homology to the gene that encodes human ribosomal L23A protein.

Clones 8-B3, P89, P98, P130 and P201 (as disclosed in U.S. Patent Application No. 09/020,956, filed February 9, 1998) were found to be contained within one
 5 contiguous sequence, referred to as P705P (SEQ ID NO: 335, with the predicted amino acid sequence provided in SEQ ID NO: 336), which was determined to be a splice variant of the known gene NKX 3.1.

Further studies on P775P resulted in the isolation of four additional sequences (SEQ ID NO: 473-476) which are all splice variants of the P775P gene. The
 10 sequence of SEQ ID NO: 474 was found to contain two open reading frames (ORFs). The predicted amino acid sequences encoded by these ORFs are provided in SEQ ID NO: 477 and 478. The cDNA sequence of SEQ ID NO: 475 was found to contain an ORF which encodes the amino acid sequence of SEQ ID NO: 479. The cDNA sequence of SEQ ID NO: 473 was found to contain four ORFs. The predicted amino acid sequences encoded by
 15 these ORFs are provided in SEQ ID NO: 480-483. Additional splice variants of P775P are provided in SEQ ID NO: 593-597.

Subsequent studies led to the identification of a genomic region on chromosome 22q11.2, known as the Cat Eye Syndrome region, that contains the five prostate genes P704P, P712P, P774P, P775P and B305D. The relative location of each of
 20 these five genes within the genomic region is shown in Fig. 10. This region may therefore be associated with malignant tumors, and other potential tumor genes may be contained within this region. These studies also led to the identification of a potential open reading frame (ORF) for P775P (provided in SEQ ID NO: 533), which encodes the amino acid sequence of SEQ ID NO: 534.

25 Comparison of the clone of SEQ ID NO: 325 (referred to as P558S) with sequences in the GenBank and GeneSeq DNA databases showed that P558S is identical to the prostate-specific transglutaminase gene, which is known to have two forms. The full-length sequences for the two forms are provided in SEQ ID NO: 773 and 774, with the corresponding amino acid sequences being provided in SEQ ID NO: 775 and 776,

respectively. The cDNA sequence of SEQ ID NO: 774 has a 15 pair base insert, resulting in a 5 amino acid insert in the corresponding amino acid sequence (SEQ ID NO: 776). This insert is not present in the sequence of SEQ ID NO: 773.

5

EXAMPLE 4

SYNTHESIS OF POLYPEPTIDES

Polypeptides may be synthesized on a Perkin Elmer/Applied Biosystems 430A peptide synthesizer using FMOC chemistry with HPTU (O-Benzotriazole-N,N,N',N'-tetramethyluronium hexafluorophosphate) activation. A Gly-Cys-Gly sequence may be attached to the amino terminus of the peptide to provide a method of conjugation, binding to an immobilized surface, or labeling of the peptide. Cleavage of the peptides from the solid support may be carried out using the following cleavage mixture: trifluoroacetic acid:ethanedithiol:thioanisole:water:phenol (40:1:2:2:3). After cleaving for 2 hours, the peptides may be precipitated in cold methyl-t-butyl-ether. The peptide pellets may then be dissolved in water containing 0.1% trifluoroacetic acid (TFA) and lyophilized prior to purification by C18 reverse phase HPLC. A gradient of 0%-60% acetonitrile (containing 0.1% TFA) in water (containing 0.1% TFA) may be used to elute the peptides. Following lyophilization of the pure fractions, the peptides may be characterized using electrospray or other types of mass spectrometry and by amino acid analysis.

20

EXAMPLE 5

FURTHER ISOLATION AND CHARACTERIZATION OF PROSTATE-SPECIFIC POLYPEPTIDES BY PCR-BASED SUBTRACTION

25

A cDNA library generated from prostate primary tumor mRNA as described above was subtracted with cDNA from normal prostate. The subtraction was performed using a PCR-based protocol (Clontech), which was modified to generate larger fragments. Within this protocol, tester and driver double stranded cDNA were separately digested with

five restriction enzymes that recognize six-nucleotide restriction sites (MluI, MscI, PvuII, SalI and StuI). This digestion resulted in an average cDNA size of 600 bp, rather than the average size of 300 bp that results from digestion with RsaI according to the Clontech protocol. This modification did not affect the subtraction efficiency. Two tester
5 populations were then created with different adapters, and the driver library remained without adapters.

The tester and driver libraries were then hybridized using excess driver cDNA. In the first hybridization step, driver was separately hybridized with each of the two tester cDNA populations. This resulted in populations of (a) unhybridized tester
10 cDNAs, (b) tester cDNAs hybridized to other tester cDNAs, (c) tester cDNAs hybridized to driver cDNAs and (d) unhybridized driver cDNAs. The two separate hybridization reactions were then combined, and rehybridized in the presence of additional denatured driver cDNA. Following this second hybridization, in addition to populations (a) through (d), a fifth population (e) was generated in which tester cDNA with one adapter hybridized
15 to tester cDNA with the second adapter. Accordingly, the second hybridization step resulted in enrichment of differentially expressed sequences which could be used as templates for PCR amplification with adaptor-specific primers.

The ends were then filled in, and PCR amplification was performed using adaptor-specific primers. Only population (e), which contained tester cDNA that did not
20 hybridize to driver cDNA, was amplified exponentially. A second PCR amplification step was then performed, to reduce background and further enrich differentially expressed sequences.

This PCR-based subtraction technique normalizes differentially expressed cDNAs so that rare transcripts that are overexpressed in prostate tumor tissue may be
25 recoverable. Such transcripts would be difficult to recover by traditional subtraction methods.

In addition to genes known to be overexpressed in prostate tumor, seventy-seven further clones were identified. Sequences of these partial cDNAs are provided in SEQ ID NO: 29 to 305. Most of these clones had no significant homology to database

sequences. Exceptions were JTPN23 (SEQ ID NO: 231; similarity to pig valosin-containing protein), JTPN30 (SEQ ID NO: 234; similarity to rat mRNA for proteasome subunit), JTPN45 (SEQ ID NO: 243; similarity to rat *norvegicus* cytosolic NADP-dependent isocitrate dehydrogenase), JTPN46 (SEQ ID NO: 244; similarity to human subclone H8 4 d4 DNA sequence), JP1D6 (SEQ ID NO: 265; similarity to *G. gallus* dynein light chain-A), JP8D6 (SEQ ID NO: 288; similarity to human BAC clone RG016J04), JP8F5 (SEQ ID NO: 289; similarity to human subclone H8 3 b5 DNA sequence), and JP8E9 (SEQ ID NO: 299; similarity to human Alu sequence).

Additional studies using the PCR-based subtraction library consisting of a prostate tumor pool subtracted against a normal prostate pool (referred to as PT-PN PCR subtraction) yielded three additional clones. Comparison of the cDNA sequences of these clones with the most recent release of GenBank revealed no significant homologies to the two clones referred to as P715P and P767P (SEQ ID NO: 312 and 314). The remaining clone was found to show some homology to the known gene KIAA0056 (SEQ ID NO: 318). Using microarray analysis to measure mRNA expression levels in various tissues, all three clones were found to be over-expressed in prostate tumors and BPH tissues. Specifically, clone P715P was over-expressed in most prostate tumors and BPH tissues by a factor of three or greater, with elevated expression seen in the majority of normal prostate samples and in fetal tissue, but negative to low expression in all other normal tissues. Clone P767P was over-expressed in several prostate tumors and BPH tissues, with moderate expression levels in half of the normal prostate samples, and background to low expression in all other normal tissues tested.

Further analysis, by microarray as described above, of the PT-PN PCR subtraction library and of a DNA subtraction library containing cDNA from prostate tumor subtracted with a pool of normal tissue cDNAs, led to the isolation of 27 additional clones (SEQ ID NO: 340-365 and 381) which were determined to be over-expressed in prostate tumor. The clones of SEQ ID NO: 341, 342, 345, 347, 348, 349, 351, 355-359, 361, 362 and 364 were also found to be expressed in normal prostate. Expression of all 26 clones in a variety of normal tissues was found to be low or undetectable, with the exception of

P544S (SEQ ID NO: 356) which was found to be expressed in small intestine. Of the 26 clones, 11 (SEQ ID NO: 340-349 and 362) were found to show some homology to previously identified sequences. No significant homologies were found to the clones of SEQ ID NO: 350, 351, 353-361, and 363-365.

5 Comparison of the sequence of SEQ ID NO: 362 with sequences in the GenBank and GeneSeq DNA databases showed that this clone (referred to as P788P) is identical to GeneSeq Accession No. X27262, which encodes a protein found in the GeneSeq protein Accession No. Y00931. The full length cDNA sequence of P788P is shown in Figure 12A (SEQ ID NO: 777), with the corresponding predicted amino acid
10 being shown in Figure 12B (SEQ ID NO: 778). Subsequently, a full-length cDNA sequence for P788P that contains polymorphisms not found in the sequence of SEQ ID NO: 779, was cloned multiple times by PCR amplification from cDNA prepared from several RNA templates from three individuals. This determined cDNA sequence of this polymorphic variant of P788P is provided in SEQ ID NO: 779, with the corresponding
15 amino acid sequence being provided in SEQ ID NO: 780. The sequence of SEQ ID NO: 780 differs from that of SEQ ID NO: 778 by six amino acid residues. The P788P protein has 7 potential transmembrane domains at the C-terminal portion and is predicted to be a plasma membrane protein with an extracellular N-terminal region.

 Further studies on the clone of SEQ ID NO: 352 (referred to as P790P) led
20 to the isolation of the full-length cDNA sequence of SEQ ID NO: 526. The corresponding predicted amino acid is provided in SEQ ID NO: 527. Data from two quantitative PCR experiments indicated that P790P is over-expressed in 11/15 tested prostate tumor samples and is expressed at low levels in spinal cord, with no expression being seen in all other normal samples tested. Data from further PCR experiments and microarray experiments
25 showed over-expression in normal prostate and prostate tumor with little or no expression in other tissues tested. P790P was subsequently found to show significant homology to a previously identified G-protein coupled prostate tissue receptor.

 Additional studies on the clone of SEQ ID NO: 354 (referred to as P776P) led to the isolation of an extended cDNA sequence, provided in SEQ ID NO: 569. The

determined cDNA sequences of three additional splice variants of P776P are provided in SEQ ID NO: 570-572. The amino acid sequences encoded by two predicted open reading frames (ORFs) contained within SEQ ID NO: 570, one predicted ORF contained within SEQ ID NO: 571, and 11 predicted ORFs contained within SEQ ID NO: 569, are provided
 5 in SEQ ID NO: 573-586, respectively.

Comparison of the cDNA sequences for the clones P767P (SEQ ID NO: 314) and P777P (SEQ ID NO: 350) with sequences in the GenBank human EST database showed that the two clones matched many EST sequences in common, suggesting that P767P and P777P may represent the same gene. A DNA consensus sequence derived
 10 from a DNA sequence alignment of P767P, P777P and multiple EST clones is provided in SEQ ID NO: 587. The amino acid sequences encoded by three putative ORFs located within SEQ ID NO: 587 are provided in SEQ ID NO: 588-590.

EXAMPLE 6

15 PEPTIDE PRIMING OF MICE AND PROPAGATION OF CTL LINES

6.1. This Example illustrates the preparation of a CTL cell line specific for cells expressing the P502S gene.

Mice expressing the transgene for human HLA A2Kb (provided by Dr L.
 20 Sherman, The Scripps Research Institute, La Jolla, CA) were immunized with P2S#12 peptide (VLGWVAEL; SEQ ID NO: 306), which is derived from the P502S gene (also referred to herein as J1-17, SEQ ID NO: 8), as described by Theobald et al., *Proc. Natl. Acad. Sci. USA* 92:11993-11997, 1995 with the following modifications. Mice were immunized with 100µg of P2S#12 and 120µg of an I-A^b binding peptide derived from
 25 hepatitis B Virus protein emulsified in incomplete Freund's adjuvant. Three weeks later these mice were sacrificed and using a nylon mesh single cell suspensions prepared. Cells were then resuspended at 6×10^6 cells/ml in complete media (RPMI-1640; Gibco BRL, Gaithersburg, MD) containing 10% FCS, 2mM Glutamine (Gibco BRL), sodium pyruvate (Gibco BRL), non-essential amino acids (Gibco BRL), 2×10^{-5} M 2-mercaptoethanol,

50U/ml penicillin and streptomycin, and cultured in the presence of irradiated (3000 rads) P2S#12-pulsed (5mg/ml P2S#12 and 10mg/ml β 2-microglobulin) LPS blasts (A2 transgenic spleens cells cultured in the presence of 7 μ g/ml dextran sulfate and 25 μ g/ml LPS for 3 days). Six days later, cells (5×10^5 /ml) were restimulated with 2.5×10^6 /ml peptide pulsed irradiated (20,000 rads) EL4A2Kb cells (Sherman et al, *Science* 258:815-818, 1992) and 3×10^6 /ml A2 transgenic spleen feeder cells. Cells were cultured in the presence of 20U/ml IL-2. Cells continued to be restimulated on a weekly basis as described, in preparation for cloning the line.

P2S#12 line was cloned by limiting dilution analysis with peptide pulsed EL4 A2Kb tumor cells (1×10^4 cells/ well) as stimulators and A2 transgenic spleen cells as feeders (5×10^5 cells/ well) grown in the presence of 30U/ml IL-2. On day 14, cells were restimulated as before. On day 21, clones that were growing were isolated and maintained in culture. Several of these clones demonstrated significantly higher reactivity (lysis) against human fibroblasts (HLA A2Kb expressing) transduced with P502S than against control fibroblasts. An example is presented in Figure 1.

This data indicates that P2S #12 represents a naturally processed epitope of the P502S protein that is expressed in the context of the human HLA A2Kb molecule.

6.2. This Example illustrates the preparation of murine CTL lines and CTL clones specific for cells expressing the P501S gene.

This series of experiments were performed similarly to that described above. Mice were immunized with the P1S#10 peptide (SEQ ID NO: 337), which is derived from the P501S gene (also referred to herein as L1-12, SEQ ID NO: 110). The P1S#10 peptide was derived by analysis of the predicted polypeptide sequence for P501S for potential HLA-A2 binding sequences as defined by published HLA-A2 binding motifs (Parker, KC, et al, *J. Immunol.*, 152:163, 1994). P1S#10 peptide was synthesized as described in Example 4, and empirically tested for HLA-A2 binding using a T cell based competition assay. Predicted A2 binding peptides were tested for their ability to compete HLA-A2 specific peptide presentation to an HLA-A2 restricted CTL clone (D150M58), which is

specific for the HLA-A2 binding influenza matrix peptide fluM58. D150M58 CTL secretes TNF in response to self-presentation of peptide fluM58. In the competition assay, test peptides at 100-200 $\mu\text{g/ml}$ were added to cultures of D150M58 CTL in order to bind HLA-A2 on the CTL. After thirty minutes, CTL cultured with test peptides, or control peptides, were tested for their antigen dose response to the fluM58 peptide in a standard TNF bioassay. As shown in Figure 3, peptide P1S#10 competes HLA-A2 restricted presentation of fluM58, demonstrating that peptide P1S#10 binds HLA-A2.

Mice expressing the transgene for human HLA A2Kb were immunized as described by Theobald et al. (*Proc. Natl. Acad. Sci. USA* 92:11993-11997, 1995) with the following modifications. Mice were immunized with 62.5 μg of P1S #10 and 120 μg of an I-A^b binding peptide derived from Hepatitis B Virus protein emulsified in incomplete Freund's adjuvant. Three weeks later these mice were sacrificed and single cell suspensions prepared using a nylon mesh. Cells were then resuspended at 6×10^6 cells/ml in complete media (as described above) and cultured in the presence of irradiated (3000 rads) P1S#10-pulsed (2 $\mu\text{g/ml}$ P1S#10 and 10mg/ml β 2-microglobulin) LPS blasts (A2 transgenic spleens cells cultured in the presence of 7 $\mu\text{g/ml}$ dextran sulfate and 25 $\mu\text{g/ml}$ LPS for 3 days). Six days later cells (5×10^5 /ml) were restimulated with 2.5×10^6 /ml peptide-pulsed irradiated (20,000 rads) EL4A2Kb cells, as described above, and 3×10^6 /ml A2 transgenic spleen feeder cells. Cells were cultured in the presence of 20 U/ml IL-2. Cells were restimulated on a weekly basis in preparation for cloning. After three rounds of *in vitro* stimulations, one line was generated that recognized P1S#10-pulsed Jurkat A2Kb targets and P501S-transduced Jurkat targets as shown in Figure 4.

A P1S#10-specific CTL line was cloned by limiting dilution analysis with peptide pulsed EL4 A2Kb tumor cells (1×10^4 cells/ well) as stimulators and A2 transgenic spleen cells as feeders (5×10^5 cells/ well) grown in the presence of 30U/ml IL-2. On day 14, cells were restimulated as before. On day 21, viable clones were isolated and maintained in culture. As shown in Figure 5, five of these clones demonstrated specific cytolytic reactivity against P501S-transduced Jurkat A2Kb targets. This data indicates that

P1S#10 represents a naturally processed epitope of the P501S protein that is expressed in the context of the human HLA-A2.1 molecule.

EXAMPLE 7

5 PRIMING OF CTL *IN VIVO* USING NAKED DNA IMMUNIZATION WITH A PROSTATE ANTIGEN

The prostate-specific antigen L1-12, as described above, is also referred to as P501S. HLA A2Kb Tg mice (provided by Dr L. Sherman, The Scripps Research Institute, La Jolla, CA) were immunized with 100 µg P501S in the vector VR1012 either
10 intramuscularly or intradermally. The mice were immunized three times, with a two week interval between immunizations. Two weeks after the last immunization, immune spleen cells were cultured with Jurkat A2Kb-P501S transduced stimulator cells. CTL lines were stimulated weekly. After two weeks of *in vitro* stimulation, CTL activity was assessed against P501S transduced targets. Two out of 8 mice developed strong anti-P501S CTL
15 responses. These results demonstrate that P501S contains at least one naturally processed HLA-A2-restricted CTL epitope.

EXAMPLE 8

20 ABILITY OF HUMAN T CELLS TO RECOGNIZE PROSTATE-SPECIFIC POLYPEPTIDES

This Example illustrates the ability of T cells specific for a prostate tumor polypeptide to recognize human tumor.

Human CD8⁺ T cells were primed *in vitro* to the P2S-12 peptide (SEQ ID NO: 306) derived from P502S (also referred to as J1-17) using dendritic cells according to
25 the protocol of Van Tsai et al. (*Critical Reviews in Immunology* 18:65-75, 1998). The resulting CD8⁺ T cell microcultures were tested for their ability to recognize the P2S-12 peptide presented by autologous fibroblasts or fibroblasts which were transduced to express the P502S gene in a γ -interferon ELISPOT assay (see Lalvani et al., *J. Exp. Med.* 186:859-865, 1997). Briefly, titrating numbers of T cells were assayed in duplicate on 10⁴

fibroblasts in the presence of 3 $\mu\text{g/ml}$ human β_2 -microglobulin and 1 $\mu\text{g/ml}$ P2S-12 peptide or control E75 peptide. In addition, T cells were simultaneously assayed on autologous fibroblasts transduced with the P502S gene or as a control, fibroblasts transduced with HER-2/*neu*. Prior to the assay, the fibroblasts were treated with 10 ng/ml γ -interferon for 5 48 hours to upregulate class I MHC expression. One of the microcultures (#5) demonstrated strong recognition of both peptide pulsed fibroblasts as well as transduced fibroblasts in a γ -interferon ELISPOT assay. Figure 2A demonstrates that there was a strong increase in the number of γ -interferon spots with increasing numbers of T cells on fibroblasts pulsed with the P2S-12 peptide (solid bars) but not with the control E75 peptide 10 (open bars). This shows the ability of these T cells to specifically recognize the P2S-12 peptide. As shown in Figure 2B, this microculture also demonstrated an increase in the number of γ -interferon spots with increasing numbers of T cells on fibroblasts transduced to express the P502S gene but not the HER-2/*neu* gene. These results provide additional confirmatory evidence that the P2S-12 peptide is a naturally processed epitope of the 15 P502S protein. Furthermore, this also demonstrates that there exists in the human T cell repertoire, high affinity T cells which are capable of recognizing this epitope. These T cells should also be capable of recognizing human tumors which express the P502S gene.

EXAMPLE 9

20 ELICITATION OF PROSTATE ANTIGEN-SPECIFIC CTL RESPONSES IN HUMAN BLOOD

This Example illustrates the ability of a prostate-specific antigen to elicit a CTL response in blood of normal humans.

25 Autologous dendritic cells (DC) were differentiated from monocyte cultures derived from PBMC of normal donors by growth for five days in RPMI medium containing 10% human serum, 50 ng/ml GMCSF and 30 ng/ml IL-4. Following culture, DC were infected overnight with recombinant P501S-expressing vaccinia virus at an M.O.I. of 5 and matured for 8 hours by the addition of 2 micrograms/ml CD40 ligand. Virus was

inactivated by UV irradiation, CD8⁺ cells were isolated by positive selection using magnetic beads, and priming cultures were initiated in 24-well plates. Following five stimulation cycles using autologous fibroblasts retrovirally transduced to express P501S and CD80, CD8⁺ lines were identified that specifically produced interferon-gamma when stimulated with autologous P501S-transduced fibroblasts. The P501S-specific activity of cell line 3A-1 could be maintained following additional stimulation cycles on autologous B-LCL transduced with P501S. Line 3A-1 was shown to specifically recognize autologous B-LCL transduced to express P501S, but not EGFP-transduced autologous B-LCL, as measured by cytotoxicity assays (⁵¹Cr release) and interferon-gamma production (Interferon-gamma Elispot; *see above and Lalvani et al., J. Exp. Med.* 186:859-865, 1997). The results of these assays are presented in Figures 6A and 6B.

EXAMPLE 10

IDENTIFICATION OF A NATURALLY PROCESSED CTL EPIOTOPE CONTAINED WITHIN A PROSTATE-SPECIFIC ANTIGEN

The 9-mer peptide p5 (SEQ ID NO: 338) was derived from the P703P antigen (also referred to as P20). The p5 peptide is immunogenic in human HLA-A2 donors and is a naturally processed epitope. Antigen specific human CD8⁺ T cells can be primed following repeated *in vitro* stimulations with monocytes pulsed with p5 peptide. These CTL specifically recognize p5-pulsed and P703P-transduced target cells in both ELISPOT (as described above) and chromium release assays. Additionally, immunization of HLA-A2Kb transgenic mice with p5 leads to the generation of CTL lines which recognize a variety of HLA-A2Kb or HLA-A2 transduced target cells expressing P703P.

Initial studies demonstrating that p5 is a naturally processed epitope were done using HLA-A2Kb transgenic mice. HLA-A2Kb transgenic mice were immunized subcutaneously in the footpad with 100 µg of p5 peptide together with 140 µg of hepatitis B virus core peptide (a Th peptide) in Freund's incomplete adjuvant. Three weeks post immunization, spleen cells from immunized mice were stimulated *in vitro* with peptide-

pulsed LPS blasts. CTL activity was assessed by chromium release assay five days after primary *in vitro* stimulation. Retrovirally transduced cells expressing the control antigen P703P and HLA-A2Kb were used as targets. CTL lines that specifically recognized both p5-pulsed targets as well as P703P-expressing targets were identified.

5 Human *in vitro* priming experiments demonstrated that the p5 peptide is immunogenic in humans. Dendritic cells (DC) were differentiated from monocyte cultures derived from PBMC of normal human donors by culturing for five days in RPMI medium containing 10% human serum, 50 ng/ml human GM-CSF and 30 ng/ml human IL-4. Following culture, the DC were pulsed with 1 ug/ml p5 peptide and cultured with CD8+ T
10 cell enriched PBMC. CTL lines were restimulated on a weekly basis with p5-pulsed monocytes. Five to six weeks after initiation of the CTL cultures, CTL recognition of p5-pulsed target cells was demonstrated. CTL were additionally shown to recognize human cells transduced to express P703P, demonstrating that p5 is a naturally processed epitope.

Studies identifying a further peptide epitope (referred to as peptide 4)
15 derived from the prostate tumor-specific antigen P703P that is capable of being recognized by CD4 T cells on the surface of cells in the context of HLA class II molecules were carried out as follows. The amino acid sequence for peptide 4 is provided in SEQ ID NO: 781, with the corresponding cDNA sequence being provided in SEQ ID NO: 782.

Twenty 15-mer peptides overlapping by 10 amino acids and derived from
20 the carboxy-terminal fragment of P703P were generated using standard procedures. Dendritic cells (DC) were derived from PBMC of a normal female donor using GM-CSF and IL-4 by standard protocols. CD4 T cells were generated from the same donor as the DC using MACS beads and negative selection. DC were pulsed overnight with pools of the 15-mer peptides, with each peptide at a final concentration of 0.25 microgram/ml.
25 Pulsed DC were washed and plated at 1×10^4 cells/well of 96-well V-bottom plates and purified CD4 T cells were added at 1×10^5 /well. Cultures were supplemented with 60 ng/ml IL-6 and 10 ng/ml IL-12 and incubated at 37 °C. Cultures were restimulated as above on a weekly basis using DC generated and pulsed as above as antigen presenting cells, supplemented with 5 ng/ml IL-7 and 10 u/ml IL-2. Following 4 *in vitro* stimulation
30 cycles, 96 lines (each line corresponding to one well) were tested for specific proliferation

and cytokine production in response to the stimulating pools with an irrelevant pool of peptides derived from mammaglobin being used as a control.

One line (referred to as 1-F9) was identified from pool #1 that demonstrated specific proliferation (measured by 3H proliferation assays) and cytokine production (measured by interferon-gamma ELISA assays) in response to pool #1 of P703P peptides. This line was further tested for specific recognition of the peptide pool, specific recognition of individual peptides in the pool, and in HLA mismatch analyses to identify the relevant restricting allele. Line 1-F9 was found to specifically proliferate and produce interferon-gamma in response to peptide pool #1, and also to peptide 4 (SEQ ID NO: 781). Peptide 4 corresponds to amino acids 126-140 of SEQ ID NO: 327. Peptide titration experiments were conducted to assess the sensitivity of line 1-F9 for the specific peptide. The line was found to specifically respond to peptide 4 at concentrations as low as 0.25 ng/ml, indicating that the T cells are very sensitive and therefore likely to have high affinity for the epitope.

To determine the HLA restriction of the P703P response, a panel of antigen presenting cells (APC) was generated that was partially matched with the donor used to generate the T cells. The APC were pulsed with the peptide and used in proliferation and cytokine assays together with line 1-F9. APC matched with the donor at HLA-DRB0701 and HLA-DQB02 alleles were able to present the peptide to the T cells, indicating that the P703P-specific response is restricted to one of these alleles.

Antibody blocking assays were utilized to determine if the restricting allele was HLA-DR0701 or HLA-DQ02. The anti-HLA-DR blocking antibody L243 or an irrelevant isotype matched IgG2a were added to T cells and APC cultures pulsed with the peptide RMPTVLQCVNVS VVS (SEQ ID NO: 781) at 250 ng/ml. Standard interferon-gamma and proliferation assays were performed. Whereas the control antibody had no effect on the ability of the T cells to recognize peptide-pulsed APC, in both assays the anti-HLA-DR antibody completely blocked the ability of the T cells to specifically recognize peptide-pulsed APC.

To determine if the peptide epitope RMPTVLQCVNVS VVS (SEQ ID NO: 781) was naturally processed, the ability of line 1-F9 to recognize APC pulsed with recombinant P703P protein was examined. For these experiments a number of recombinant

P703P sources were utilized; E. coli-derived P703P, Pichia-derived P703P and baculovirus-derived P703P. Irrelevant protein controls used were E. coli derived-L3E and baculovirus-derived mammaglobin. In interferon-gamma ELISA assays, line 1-F9 was able to efficiently recognize both E. coli forms of P703P as well as Pichia-derived
 5 recombinant P703P, while baculovirus-derived P703P was recognized less efficiently. Subsequent Western blot analysis revealed that the E coli and Pichia P703P protein preparations were intact while the baculovirus P703P preparation was approximately 75% degraded. Thus, peptide RMPTVLQCVNVSVVS (SEQ ID NO: 781) from P703P is a naturally processed peptide epitope derived from P703P and presented to T cells in the
 10 context of HLA-DRB-0701

In further studies, twenty-four 15-mer peptides overlapping by 10 amino acids and derived from the N-terminal fragment of P703P (corresponding to amino acids 27-154 of SEQ ID NO: 525) were generated by standard procedures and their ability to be recognized by CD4 cells was determined essentially as described above. DC were pulsed
 15 overnight with pools of the peptides with each peptide at a final concentration of 10 microgram/ml. A large number of individual CD4 T cell lines (65/480) demonstrated significant proliferation and cytokine release (IFN-gamma) in response to the P703P peptide pools but not to a control peptide pool. The CD4 T cell lines which demonstrated specific activity were restimulated on the appropriate pool of P703P peptides and reassayed
 20 on the individual peptides of each pool as well as a peptide dose titration of the pool of peptides in a IFN-gamma release assay and in a proliferation assay.

Sixteen immunogenic peptides were recognized by the T cells from the entire set of peptide antigens tested. The amino acid sequences of these peptides are provided in SEQ ID NO: 799-814, with the corresponding cDNA sequences being
 25 provided in SEQ ID NO: 783-798, respectively. In some cases the peptide reactivity of the T cell line could be mapped to a single peptide, however some could be mapped to more than one peptide in each pool. Those CD4 T cell lines that displayed a representative pattern of recognition from each peptide pool with a reasonable affinity for peptide were chosen for further analysis (I-1A, -6A; II-4C, -5E; III-6E, IV-4B, -3F, -9B, -10F, V-5B, -
 30 4D, and -10F). These CD4 T cells lines were restimulated on the appropriate individual

peptide and reassayed on autologous DC pulsed with a truncated form of recombinant P703P protein made in *E. coli* (a.a. 96 - 254 of SEQ ID NO: 525), full-length P703P made in the baculovirus expression system, and a fusion between influenza virus NS1 and P703P made in *E. coli*. Of the T cell lines tested, line I-1A recognized specifically the truncated form of P703P (*E. coli*) but no other recombinant form of P703P. This line also recognized the peptide used to elicit the T cells. Line 2-4C recognized the truncated form of P703P (*E. coli*) and the full length form of P703P made in baculovirus, as well as peptide. The remaining T cell lines tested were either peptide-specific only (II-5E, II-6F, IV-4B, IV-3F, IV-9B, IV-10F, V-5B and V-4D) or were non-responsive to any antigen tested (V-10F). These results demonstrate that the peptide sequence RPLLANDLMLIKLDE (SEQ ID NO: 814; corresponding to a.a. 110-124 of SEQ ID NO: 525) recognized by the T cell line I-1A, and the peptide sequences SVSESDTIRSISIAS (SEQ ID NO: 811; corresponding to a.a. 125-139 of SEQ ID NO: 525) and ISIASQCPTAGNSCL (SEQ ID NO: 810; corresponding to a.a. 135-149 of SEQ ID NO: 525) recognized by the T cell line II-4C may be naturally processed epitopes of the P703P protein.

EXAMPLE 11

EXPRESSION OF A BREAST TUMOR-DERIVED ANTIGEN IN PROSTATE

Isolation of the antigen B305D from breast tumor by differential display is described in US Patent Application No. 08/700,014, filed August 20, 1996. Several different splice forms of this antigen were isolated. The determined cDNA sequences for these splice forms are provided in SEQ ID NO: 366-375, with the predicted amino acid sequences corresponding to the sequences of SEQ ID NO: 292, 298 and 301-303 being provided in SEQ ID NO: 299-306, respectively. In further studies, a splice variant of the cDNA sequence of SEQ ID NO: 366 was isolated which was found to contain an additional guanine residue at position 884 (SEQ ID NO: 530), leading to a frameshift in the open reading frame. The determined DNA sequence of this ORF is provided in SEQ ID NO:

531. This frameshift generates a protein sequence (provided in SEQ ID NO: 532) of 293 amino acids that contains the C-terminal domain common to the other isoforms of B305D but that differs in the N-terminal region.

The expression levels of B305D in a variety of tumor and normal tissues were examined by real time PCR and by Northern analysis. The results indicated that B305D is highly expressed in breast tumor, prostate tumor, normal prostate and normal testes, with expression being low or undetectable in all other tissues examined (colon tumor, lung tumor, ovary tumor, and normal bone marrow, colon, kidney, liver, lung, ovary, skin, small intestine, stomach). Using real-time PCR on a panel of prostate tumors, expression of B305D in prostate tumors was shown to increase with increasing Gleason grade, demonstrating that expression of B305D increases as prostate cancer progresses.

EXAMPLE 12

GENERATION OF HUMAN CTL *IN VITRO* USING WHOLE GENE PRIMING AND STIMULATION

TECHNIQUES WITH PROSTATE-SPECIFIC ANTIGEN

Using *in vitro* whole-gene priming with P501S-vaccinia infected DC (see, for example, Yee et al, *The Journal of Immunology*, 157(9):4079-86, 1996), human CTL lines were derived that specifically recognize autologous fibroblasts transduced with P501S (also known as L1-12), as determined by interferon- γ ELISPOT analysis as described above. Using a panel of HLA-mismatched B-LCL lines transduced with P501S, these CTL lines were shown to be likely restricted to HLAB class I allele. Specifically, dendritic cells (DC) were differentiated from monocyte cultures derived from PBMC of normal human donors by growing for five days in RPMI medium containing 10% human serum, 50 ng/ml human GM-CSF and 30 ng/ml human IL-4. Following culture, DC were infected overnight with recombinant P501S vaccinia virus at a multiplicity of infection (M.O.I) of five, and matured overnight by the addition of 3 μ g/ml CD40 ligand. Virus was inactivated by UV irradiation. CD8⁺ T cells were isolated using a magnetic bead system, and priming cultures were initiated using standard culture techniques. Cultures were restimulated every

7-10 days using autologous primary fibroblasts retrovirally transduced with P501S and CD80. Following four stimulation cycles, CD8⁺ T cell lines were identified that specifically produced interferon- γ when stimulated with P501S and CD80-transduced autologous fibroblasts. A panel of HLA-mismatched B-LCL lines transduced with P501S
 5 were generated to define the restriction allele of the response. By measuring interferon- γ in an ELISPOT assay, the P501S specific response was shown to be likely restricted by HLA B alleles. These results demonstrate that a CD8⁺ CTL response to P501S can be elicited.

To identify the epitope(s) recognized, cDNA encoding P501S was fragmented by various restriction digests, and sub-cloned into the retroviral expression
 10 vector pBIB-KS. Retroviral supernatants were generated by transfection of the helper packaging line Phoenix-Ampho. Supernatants were then used to transduce Jurkat/A2Kb cells for CTL screening. CTL were screened in IFN- γ ELISPOT assays against these A2Kb targets transduced with the “library” of P501S fragments. Initial positive fragments P501S/H3 and P501S/F2 were sequenced and found to encode amino acids 106-553 and
 15 amino acids 136-547, respectively, of SEQ ID NO: 113. A truncation of H3 was made to encode amino acid residues 106-351 of SEQ ID NO: 113, which was unable to stimulate the CTL, thus localizing the epitope to amino acid residues 351-547. Additional fragments encoding amino acids 1-472 (Fragment A) and amino acids 1-351 (Fragment B) were also constructed. Fragment A but not Fragment B stimulated the CTL thus localizing the
 20 epitope to amino acid residues 351-472. Overlapping 20-mer and 18-mer peptides representing this region were tested by pulsing Jurkat/A2Kb cells versus CTL in an IFN- γ assay. Only peptides P501S-369(20) and P501S-369(18) stimulated the CTL. Nine-mer and 10-mer peptides representing this region were synthesized and similarly tested. Peptide P501S-370 (SEQ ID NO: 539) was the minimal 9-mer giving a strong
 25 response. Peptide P501S-376 (SEQ ID NO: 540) also gave a weak response, suggesting that it might represent a cross-reactive epitope.

In subsequent studies, the ability of primary human B cells transduced with P501S to prime MHC class I-restricted, P501S-specific, autologous CD8 T cells was examined. Primary B cells were derived from PBMC of a homozygous HLA-A2 donor by

culture in CD40 ligand and IL-4, transduced at high frequency with recombinant P501S in the vector pBIB, and selected with blastocidin-S. For *in vitro* priming, purified CD8⁺ T cells were cultured with autologous CD40 ligand + IL-4 derived, P501S-transduced B cells in a 96-well microculture format. These CTL microcultures were re-stimulated with

5 P501S-transduced B cells and then assayed for specificity. Following this initial screen, microcultures with significant signal above background were cloned on autologous EBV-transformed B cells (BLCL), also transduced with P501S. Using IFN-gamma ELISPOT for detection, several of these CD8 T cell clones were found to be specific for P501S, as demonstrated by reactivity to BLCL/P501S but not BLCL transduced with control antigen.

10 It was further demonstrated that the anti-P501S CD8 T cell specificity is HLA-A2-restricted. First, antibody blocking experiments with anti-HLA-A,B,C monoclonal antibody (W6.32), anti-HLA-B,C monoclonal antibody (B1.23.2) and a control monoclonal antibody showed that only the anti-HLA-A,B,C antibody blocked recognition of P501S-expressing autologous BLCL. Secondly, the anti-P501S CTL also recognized an HLA-A2

15 matched, heterologous BLCL transduced with P501S, but not the corresponding EGFP transduced control BLCL.

EXAMPLE 13

IDENTIFICATION OF PROSTATE-SPECIFIC ANTIGENS

20 BY MICROARRAY ANALYSIS

This Example describes the isolation of certain prostate-specific polypeptides from a prostate tumor cDNA library.

A human prostate tumor cDNA expression library as described above was

25 screened using microarray analysis to identify clones that display at least a three fold over-expression in prostate tumor and/or normal prostate tissue, as compared to non-prostate normal tissues (not including testis). 372 clones were identified, and 319 were successfully sequenced. Table I presents a summary of these clones, which are shown in SEQ ID NOs:385-400. Of these sequences SEQ ID NOs:386, 389, 390 and 392 correspond to

novel genes, and SEQ ID NOs: 393 and 396 correspond to previously identified sequences. The others (SEQ ID NOs:385, 387, 388, 391, 394, 395 and 397-400) correspond to known sequences, as shown in Table I.

Table I

002690-032990

Summary of Prostate Tumor Antigens

Known Genes	Previously Identified Genes	Novel Genes
T-cell gamma chain	P504S	23379 (SEQ ID NO:389)
Kallikrein	P1000C	23399 (SEQ ID NO:392)
Vector	P501S	23320 (SEQ ID NO:386)
CGI-82 protein mRNA (23319; SEQ ID NO:385)	P503S	23381 (SEQ ID NO:390)
PSA	P510S	
Ald. 6 Dehyd.	P784P	
L-idoitol-2 dehydrogenase (23376; SEQ ID NO:388)	P502S	
Ets transcription factor PDEF (22672; SEQ ID NO:398)	P706P	
hTGR (22678; SEQ ID NO:399)	19142.2, bangur.seq (22621; SEQ ID NO:396)	
KIAA0295(22685; SEQ ID NO:400)	5566.1 Wang (23404; SEQ ID NO:393)	
Prostatic Acid Phosphatase(22655; SEQ ID NO:397)	P712P	
transglutaminase (22611; SEQ ID NO:395)	P778P	
HDLBP (23508; SEQ ID NO:394)		
CGI-69 Protein(23367; SEQ ID NO:387)		
KIAA0122(23383; SEQ ID NO:391)		
TEEG		

CGI-82 showed 4.06 fold over-expression in prostate tissues as compared to
5 other normal tissues tested. It was over-expressed in 43% of prostate tumors, 25% normal

prostate, not detected in other normal tissues tested. L-iditol-2 dehydrogenase showed 4.94 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 90% of prostate tumors, 100% of normal prostate, and not detected in other normal tissues tested. Ets transcription factor PDEF showed 5.55 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 47% prostate tumors, 25% normal prostate and not detected in other normal tissues tested. hTGR1 showed 9.11 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 63% of prostate tumors and is not detected in normal tissues tested including normal prostate. KIAA0295 showed 5.59 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 47% of prostate tumors, low to undetectable in normal tissues tested including normal prostate tissues. Prostatic acid phosphatase showed 9.14 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 67% of prostate tumors, 50% of normal prostate, and not detected in other normal tissues tested. Transglutaminase showed 14.84 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 30% of prostate tumors, 50% of normal prostate, and is not detected in other normal tissues tested. High density lipoprotein binding protein (HDLBP) showed 28.06 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 97% of prostate tumors, 75% of normal prostate, and is undetectable in all other normal tissues tested. CGI-69 showed 3.56 fold over-expression in prostate tissues as compared to other normal tissues tested. It is a low abundant gene, detected in more than 90% of prostate tumors, and in 75% normal prostate tissues. The expression of this gene in normal tissues was very low. KIAA0122 showed 4.24 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 57% of prostate tumors, it was undetectable in all normal tissues tested including normal prostate tissues. 19142.2 bangur showed 23.25 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 97% of prostate tumors and 100% of normal prostate. It was undetectable in other normal tissues tested. 5566.1 Wang showed 3.31 fold over-

expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 97% of prostate tumors, 75% normal prostate and was also over-expressed in normal bone marrow, pancreas, and activated PBMC. Novel clone 23379 (also referred to as P553S) showed 4.86 fold over-expression in prostate tissues as compared to other normal tissues tested. It was detectable in 97% of prostate tumors and 75% normal prostate and is undetectable in all other normal tissues tested. Novel clone 23399 showed 4.09 fold over-expression in prostate tissues as compared to other normal tissues tested. It was over-expressed in 27% of prostate tumors and was undetectable in all normal tissues tested including normal prostate tissues. Novel clone 23320 showed 3.15 fold over-expression in prostate tissues as compared to other normal tissues tested. It was detectable in all prostate tumors and 50% of normal prostate tissues. It was also expressed in normal colon and trachea. Other normal tissues do not express this gene at high level.

Subsequent full-length cloning studies on P553S, using standard techniques, revealed that this clone is an incomplete spliced form of P501S. The determined cDNA sequences for four splice variants of P553S are provided in SEQ ID NO: 702-705. An amino acid sequence encoded by SEQ ID NO: 705 is provided in SEQ ID NO: 706. The cDNA sequence of SEQ ID NO: 702 was found to contain two open reading frames (ORFs). The amino acid sequences encoded by these two ORFs are provided in SEQ ID NO: 707 and 708.

EXAMPLE 14

IDENTIFICATION OF PROSTATE-SPECIFIC ANTIGENS

BY ELECTRONIC SUBTRACTION

This Example describes the use of an electronic subtraction technique to identify prostate-specific antigens.

Potential prostate-specific genes present in the GenBank human EST database were identified by electronic subtraction (similar to that described by Vasmatizis et al., *Proc. Natl. Acad. Sci. USA* 95:300-304, 1998). The sequences of EST clones

(43,482) derived from various prostate libraries were obtained from the GenBank public human EST database. Each prostate EST sequence was used as a query sequence in a BLASTN (National Center for Biotechnology Information) search against the human EST database. All matches considered identical (length of matching sequence >100 base pairs, density of identical matches over this region > 70%) were grouped (aligned) together in a cluster. Clusters containing more than 200 ESTs were discarded since they probably represented repetitive elements or highly expressed genes such as those for ribosomal proteins. If two or more clusters shared common ESTs, those clusters were grouped together into a "supercluster," resulting in 4,345 prostate superclusters.

Records for the 479 human cDNA libraries represented in the GenBank release were downloaded to create a database of these cDNA library records. These 479 cDNA libraries were grouped into three groups: Plus (normal prostate and prostate tumor libraries, and breast cell line libraries, in which expression was desired), Minus (libraries from other normal adult tissues, in which expression was not desirable), and Other (libraries from fetal tissue, infant tissue, tissues found only in women, non-prostate tumors and cell lines other than prostate cell lines, in which expression was considered to be irrelevant). A summary of these library groups is presented in Table II.

Table II

Prostate cDNA Libraries and ESTs

Library	# of Libraries	# of ESTs
Plus	25	43,482
Normal	11	18,875
Tumor	11	21,769
Cell lines	3	2,838
Minus	166	
Other	287	

Each supercluster was analyzed in terms of the ESTs within the supercluster. The tissue source of each EST clone was noted and used to classify the superclusters into four groups: Type 1- EST clones found in the Plus group libraries only; no expression detected in Minus or Other group libraries; Type 2- EST clones derived from the Plus and Other group libraries only; no expression detected in the Minus group; Type 3- EST clones derived from the Plus, Minus and Other group libraries, but the number of ESTs derived from the Plus group is higher than in either the Minus or Other groups; and Type 4- EST clones derived from Plus, Minus and Other group libraries, but the number derived from the Plus group is higher than the number derived from the Minus group. This analysis identified 4,345 breast clusters (*see* Table III). From these clusters, 3,172 EST clones were ordered from Research Genetics, Inc., and were received as frozen glycerol stocks in 96-well plates.

Table III

Prostate Cluster Summary

Type	# of Superclusters	# of ESTs Ordered
1	688	677
2	2899	2484
3	85	11
4	673	0
Total	4345	3172

The EST clone inserts were PCR-amplified using amino-linked PCR primers for Synteni microarray analysis. When more than one PCR product was obtained for a particular clone, that PCR product was not used for expression analysis. In total, 2,528 clones from the electronic subtraction method were analyzed by microarray analysis to identify electronic subtraction breast clones that had high levels of tumor vs. normal

tissue mRNA. Such screens were performed using a Synteni (Palo Alto, CA) microarray, according to the manufacturer's instructions (and essentially as described by Schena et al., *Proc. Natl. Acad. Sci. USA* 93:10614-10619, 1996 and Heller et al., *Proc. Natl. Acad. Sci. USA* 94:2150-2155, 1997). Within these analyses, the clones were arrayed on the chip, which was then probed with fluorescent probes generated from normal and tumor prostate cDNA, as well as various other normal tissues. The slides were scanned and the fluorescence intensity was measured.

Clones with an expression ratio greater than 3 (*i.e.*, the level in prostate tumor and normal prostate mRNA was at least three times the level in other normal tissue mRNA) were identified as prostate tumor-specific sequences (Table IV). The sequences of these clones are provided in SEQ ID NO: 401-453, with certain novel sequences shown in SEQ ID NO: 407, 413, 416-419, 422, 426, 427 and 450.

Table IV

Prostate-tumor Specific Clones

SEQ ID NO.	Sequence Designation	Comments
401	22545	previously identified P1000C
402	22547	previously identified P704P
403	22548	known
404	22550	known
405	22551	PSA
406	22552	prostate secretory protein 94
407	22553	novel
408	22558	previously identified P509S
409	22562	glandular kallikrein
410	22565	previously identified P1000C
411	22567	PAP
412	22568	B1006C (breast tumor antigen)
413	22570	novel
414	22571	PSA
415	22572	previously identified P706P
416	22573	novel
417	22574	novel

418	22575	novel
419	22580	novel
420	22581	PAP
421	22582	prostatic secretory protein 94
422	22583	novel
423	22584	prostatic secretory protein 94
424	22585	prostatic secretory protein 94
425	22586	known
426	22587	novel
427	22588	novel
428	22589	PAP
429	22590	known
430	22591	PSA
431	22592	known
432	22593	Previously identified P777P
433	22594	T cell receptor gamma chain
434	22595	Previously identified P705P
435	22596	Previously identified P707P
436	22847	PAP
437	22848	known
438	22849	prostatic secretory protein 57
439	22851	PAP
440	22852	PAP
441	22853	PAP
442	22854	previously identified P509S
443	22855	previously identified P705P
444	22856	previously identified P774P
445	22857	PSA
446	23601	previously identified P777P
447	23602	PSA
448	23605	PSA
449	23606	PSA
450	23612	novel
451	23614	PSA
452	23618	previously identified P1000C
453	23622	previously identified P705P

Further studies on the clone of SEQ ID NO: 407 (also referred to as P1020C) led to the isolation of an extended cDNA sequence provided in SEQ ID NO: 591. This extended cDNA sequence was found to contain an open reading frame that encodes

the predicted amino acid sequence of SEQ ID NO: 592. The P1020C cDNA and amino acid sequences were found to show some similarity to the human endogenous retroviral HERV-K pol gene and protein.

5

EXAMPLE 15

FURTHER IDENTIFICATION OF PROSTATE-SPECIFIC ANTIGENS BY MICROARRAY ANALYSIS

This Example describes the isolation of additional prostate-specific polypeptides from a prostate tumor cDNA library.

10

A human prostate tumor cDNA expression library as described above was screened using microarray analysis to identify clones that display at least a three fold over-expression in prostate tumor and/or normal prostate tissue, as compared to non-prostate normal tissues (not including testis). 142 clones were identified and sequenced. Certain of these clones are shown in SEQ ID NO: 454-467. Of these sequences, SEQ ID NO: 459-15 461 represent novel genes. The others (SEQ ID NO: 454-458 and 461-467) correspond to known sequences.

EXAMPLE 16

FURTHER CHARACTERIZATION OF PROSTATE-SPECIFIC ANTIGEN P710P

20

This Example describes the full length cloning of P710P.

The prostate cDNA library described above was screened with the P710P fragment described above. One million colonies were plated on LB/Ampicillin plates. Nylon membrane filters were used to lift these colonies, and the cDNAs picked up by these 25 filters were then denatured and cross-linked to the filters by UV light. The P710P fragment was radiolabeled and used to hybridize with the filters. Positive cDNA clones were selected and their cDNAs recovered and sequenced by an automatic Perkin Elmer/Applied Biosystems Division Sequencer. Four sequences were obtained, and are presented in SEQ ID NO: 468-471. These sequences appear to represent different splice variants of the

P710P gene. Subsequent comparison of the cDNA sequences of P710P with those in Genbank revealed homology to the DD3 gene (Genbank accession numbers AF103907 & AF103908). The cDNA sequence of DD3 is provided in SEQ ID NO: 690.

5

EXAMPLE 17

PROTEIN EXPRESSION OF PROSTATE-SPECIFIC ANTIGENS

This example describes the expression and purification of prostate-specific antigens in *E. coli*, baculovirus and mammalian cells.

10 **A) EXPRESSION OF P501S IN *E. COLI***

Expression of the full-length form of P501S was attempted by first cloning P501S without the leader sequence (amino acids 36-553 of SEQ ID NO: 113) downstream of the first 30 amino acids of the *M. tuberculosis* antigen Ra12 (SEQ ID NO: 484) in pET17b. Specifically, P501S DNA was used to perform PCR using the primers AW025 (SEQ ID NO: 485) and AW003 (SEQ ID NO: 486). AW025 is a sense cloning primer that contains a HindIII site. AW003 is an antisense cloning primer that contains an EcoRI site. DNA amplification was performed using 5 µl 10X Pfu buffer, 1 µl 20 mM dNTPs, 1 µl each of the PCR primers at 10 µM concentration, 40 µl water, 1 µl Pfu DNA polymerase (Stratagene, La Jolla, CA) and 1 µl DNA at 100 ng/µl. Denaturation at 95°C was performed for 30 sec, followed by 10 cycles of 95°C for 30 sec, 60°C for 1 min and by 72°C for 3 min. 20 cycles of 95°C for 30 sec, 65°C for 1 min and by 72°C for 3 min, and lastly by 1 cycle of 72°C for 10 min. The PCR product was cloned to Ra12m/pET17b using HindIII and EcoRI. The sequence of the resulting fusion construct (referred to as Ra12-P501S-F) was confirmed by DNA sequencing.

25 The fusion construct was transformed into BL21(DE3)pLysE, pLysS and CodonPlus *E. coli* (Stratagene) and grown overnight in LB broth with kanamycin. The resulting culture was induced with IPTG. Protein was transferred to PVDF membrane and blocked with 5% non-fat milk (in PBS-Tween buffer), washed three times and incubated

with mouse anti-His tag antibody (Clontech) for 1 hour. The membrane was washed 3 times and probed with HRP-Protein A (Zymed) for 30 min. Finally, the membrane was washed 3 times and developed with ECL (Amersham). No expression was detected by Western blot. Similarly, no expression was detected by Western blot when the Ra12-
 5 P501S-F fusion was used for expression in BL21CodonPlus by CE6 phage (Invitrogen).

An N-terminal fragment of P501S (amino acids 36-325 of SEQ ID NO: 113) was cloned down-stream of the first 30 amino acids of the *M. tuberculosis* antigen Ra12 in pET17b as follows. P501S DNA was used to perform PCR using the primers AW025 (SEQ ID NO: 485) and AW027 (SEQ ID NO: 487). AW027 is an antisense cloning primer that
 10 contains an EcoRI site and a stop codon. DNA amplification was performed essentially as described above. The resulting PCR product was cloned to Ra12 in pET17b at the HindIII and EcoRI sites. The fusion construct (referred to as Ra12-P501S-N) was confirmed by DNA sequencing.

The Ra12-P501S-N fusion construct was used for expression in
 15 BL21(DE3)pLysE, pLysS and CodonPlus, essentially as described above. Using Western blot analysis, protein bands were observed at the expected molecular weight of 36 kDa. Some high molecular weight bands were also observed, probably due to aggregation of the recombinant protein. No expression was detected by Western blot when the Ra12-P501S-F fusion was used for expression in BL21CodonPlus by CE6 phage.

A fusion construct comprising a C-terminal portion of P501S (amino acids 257-553 of SEQ ID NO: 113) located down-stream of the first 30 amino acids of the *M. tuberculosis* antigen Ra12 (SEQ ID NO: 484) was prepared as follows. P501S DNA was used to perform PCR using the primers AW026 (SEQ ID NO: 488) and AW003 (SEQ ID NO: 486). AW026 is a sense cloning primer that contains a HindIII site. DNA
 25 amplification was performed essentially as described above. The resulting PCR product was cloned to Ra12 in pET17b at the HindIII and EcoRI sites. The sequence for the fusion construct (referred to as Ra12-P501S-C) was confirmed.

The Ra12-P501S-C fusion construct was used for expression in BL21(DE3)pLysE, pLysS and CodonPlus, as described above. A small amount of protein

was detected by Western blot, with some molecular weight aggregates also being observed. Expression was also detected by Western blot when the Ra12-P501S-C fusion was used for expression in BL21CodonPlus induced by CE6 phage.

B) EXPRESSION OF P501S IN BACULOVIRUS

5 The Bac-to-Bac baculovirus expression system (BRL Life Technologies, Inc.) was used to express P501S protein in insect cells. Full-length P501S (SEQ ID NO: 113) was amplified by PCR and cloned into the XbaI site of the donor plasmid pFastBacI. The recombinant bacmid and baculovirus were prepared according to the manufacturer's instructions. The recombinant baculovirus was amplified in Sf9 cells and the high titer
10 viral stocks were utilized to infect High Five cells (Invitrogen) to make the recombinant protein. The identity of the full-length protein was confirmed by N-terminal sequencing of the recombinant protein and by Western blot analysis (Figure 7). Specifically, 0.6 million High Five cells in 6-well plates were infected with either the unrelated control virus BV/ECD_PD (lane 2), with recombinant baculovirus for P501S at different amounts or
15 MOIs (lanes 4-8), or were uninfected (lane 3). Cell lysates were run on SDS-PAGE under reducing conditions and analyzed by Western blot with the anti-P501S monoclonal antibody P501S-10E3-G4D3 (prepared as described below). Lane 1 is the biotinylated protein molecular weight marker (BioLabs).

 The localization of recombinant P501S in the insect cells was investigated
20 as follows. The insect cells overexpressing P501S were fractionated into fractions of nucleus, mitochondria, membrane and cytosol. Equal amounts of protein from each fraction were analyzed by Western blot with a monoclonal antibody against P501S. Due to the scheme of fractionation, both nucleus and mitochondria fractions contain some plasma membrane components. However, the membrane fraction is basically free from
25 mitochondria and nucleus. P501S was found to be present in all fractions that contain the membrane component, suggesting that P501S may be associated with plasma membrane of the insect cells expressing the recombinant protein.

C) EXPRESSION OF P501S IN MAMMALIAN CELLS

Full-length P501S (553AA) was cloned into various mammalian expression vectors, including pCEP4 (Invitrogen), pVR1012 (Vical, San Diego, CA) and a modified form of the retroviral vector pBMN, referred to as pBIB. Transfection of P501S/pCEP4 and P501S/pVR1012 into HEK293 fibroblasts was carried out using the Fugene transfection reagent (Boehringer Mannheim). Briefly, 2 μ l of Fugene reagent was diluted into 100 μ l of serum-free media and incubated at room temperature for 5-10 min. This mixture was added to 1 μ g of P501S plasmid DNA, mixed briefly and incubated for 30 minutes at room temperature. The Fugene/DNA mixture was added to cells and incubated for 24-48 hours. Expression of recombinant P501S in transfected HEK293 fibroblasts was detected by means of Western blot employing a monoclonal antibody to P501S.

Transfection of p501S/pCEP4 into CHO-K cells (American Type Culture Collection, Rockville, MD) was carried out using GenePorter transfection reagent (Gene Therapy Systems, San Diego, CA). Briefly, 15 μ l of GenePorter was diluted in 500 μ l of serum-free media and incubated at room temperature for 10 min. The GenePorter/media mixture was added to 2 μ g of plasmid DNA that was diluted in 500 μ l of serum-free media, mixed briefly and incubated for 30 min at room temperature. CHO-K cells were rinsed in PBS to remove serum proteins, and the GenePorter/DNA mix was added and incubated for 5 hours. The transfected cells were then fed an equal volume of 2x media and incubated for 24-48 hours.

FACS analysis of P501S transiently infected CHO-K cells, demonstrated surface expression of P501S. Expression was detected using rabbit polyclonal antisera raised against a P501S peptide, as described below. Flow cytometric analysis was performed using a FaCScan (Becton Dickinson), and the data were analyzed using the Cell Quest program.

D) EXPRESSION OF P703P IN BACULOVIRUS

The cDNA for full-length P703P-DE5 (SEQ ID NO: 326), together with several flanking restriction sites, was obtained by digesting the plasmid pCDNA703 with

restriction endonucleases Xba I and Hind III. The resulting restriction fragment (approx. 800 base pairs) was ligated into the transfer plasmid pFastBacI which was digested with the same restriction enzymes. The sequence of the insert was confirmed by DNA sequencing. The recombinant transfer plasmid pFBP703 was used to make recombinant bacmid DNA and baculovirus using the Bac-To-Bac Baculovirus expression system (BRL Life Technologies). High Five cells were infected with the recombinant virus BVP703, as described above, to obtain recombinant P703P protein.

E) EXPRESSION OF P788P IN *E. COLI*

A truncated, N-terminal portion, of P788P (residues 1-644 of SEQ ID NO: 777; referred to as P788P-N) fused with a C-terminal 6xHis Tag was expressed in *E. coli* as follows. P788P cDNA was amplified using the primers AW080 and AW081 (SEQ ID NO: 815 and 816). AW080 is a sense cloning primer with an NdeI site. AW081 is an antisense cloning primer with a XhoI site. The PCR-amplified P788P, as well as the vector pCRX1, were digested with NdeI and XhoI. Vector and insert were ligated and transformed into NovaBlue cells. Colonies were randomly screened for insert and then sequenced. P788P-N clone #6 was confirmed to be identical to the designed construct. The expression construct P788P-N #6/pCRX1 was transformed into *E. coli* BL21 CodonPlus-RIL competent cells. After induction, most of the cells grew well, achieving OD600 of greater than 2.0 after 3 hr. Coomassie stained SDS-PAGE showed an over-expressed band at about 75 kD. Western blot analysis using a 6xHisTag antibody confirmed the band was P788P-N. The determined cDNA sequence for P788P-N is provided in SEQ ID NO: 817, with the corresponding amino acid sequence being provided in SEQ ID NO: 818.

F) EXPRESSION OF P510S IN *E. COLI*

The P510S protein has 9 potential transmembrane domains and is predicted to be located at the plasma membrane. The C-terminal protein of this protein, as well as the predicted third extracellular domain of P510S were expressed in *E. coli* as follows.

The expression construct referred to as Ra12-P501S-C was designed to have a 6 HisTag at the N-terminal end, followed by the *M. tuberculosis* antigen Ra12 (SEQ ID NO: 819) and then the C-terminal portion of P510S (amino residues 1176-1261 of SEQ ID NO: 538). Full-length P510S was used to amplify the P510S-C fragment by PCR using the primers AW056 and AW057 (SEQ ID NO: 820 and 821, respectively). AW056 is a sense cloning primer with an EcoRI site. AW057 is an antisense primer with stop and XhoI sites. The amplified P501S fragment and Ra12/pCRX1 were digested with EcoRI and XhoI and then purified. The insert and vector were ligated together and transformed into NovaBlue. Colonies were randomly screened for insert and sequences. For protein expression, the expression construct was transformed into *E. coli* BL21 (DE3) CodonPlus-RIL competent cells. A mini-induction screen was performed to optimize the expression conditions. After induction the cells grew well, achieving OD 600 nm greater than 2.0 after 3 hours. Coomassie stain SDS-PAGE showed a highly over-expressed band at approx. 30 kD. Though this is higher than the expected molecular weight, western blot analysis was positive, showing this band to be the His tag-containing protein. The optimized culture conditions are as follows. Dilute overnight culture/daytime culture (LB + kanamycin + chloramphenicol) into 2xYT (with kanamycin and chloramphenicol) at a ratio of 25 ml culture to 1 liter 2xYT. Allow to grow at 37 °C until OD600 = 0.6. Take an aliquot out as T0 sample. Add 1 mM IPTG and allow to grow at 30 °C for 3 hours. Take out a T3 sample, spin down cells and store at -80 °C. The determined cDNA and amino acid sequences for the Ra12-P510S-C construct are provided in SEQ ID NO: 822 and 825, respectively.

The expression construct P510S-C was designed to have a 5' added start codon and a glycine (GGA) codon and then the P510S C terminal fragment followed by the in frame 6x histidine tag and stop codon from the pET28b vector. The cloning strategy is similar to that used for Ra12-P510S-C, except that the PCR primers employed were those shown in SEQ ID NO: 828 and 829, respectively and the NcoI/XhoI cut in pET28b was used. The primer of SEQ ID NO: 828 created a 5' NcoI site and added a start codon. The antisense primer of SEQ ID NO: 829 creates a XhoI site on P510S C terminal fragment. Clones

were confirmed by sequencing. For protein expression, the expression construct was transformed into *E. coli* BL21 (DE3) CodonPlus-RIL competent cells. An OD600 of greater than 2.0 was obtained 30 hours after induction. Coomassie stained SDS-PAGE showed an over-expressed band at about 11 kD. Western blot analysis confirmed that the

5 band was P510S-C, as did N-terminal protein sequencing. The optimized culture conditions are as follows: dilute overnight culture/daytime culture (LB + kanamycin + chloramphenicol) into 2x YT (+ kanamycin and chloramphenicol) at a ratio of 25 mL culture to 1 liter 2x YT, and allow to grow at 37 °C until an OD 600 of about 0.5 is reached. Take out an aliquot as T0 sample. Add 1 mM IPTG and allow to grow at 30 °C

10 for 3 hours. Spin down the cells and store at -80 °C until purification. The determined cDNA and amino acid sequence for the P510S-C construct are shown in SEQ ID NO: 823 and 826, respectively.

The predicted third extracellular domain of P510S (P510S-E3; residues 328-676 of SEQ ID NO: 538) was expressed in *E. coli* as follows. The P510S fragment was

15 amplified by PCR using the primers shown in SEQ ID NO: 830 and 831. The primer of SEQ ID NO: 830 is a sense primer with an NdeI site for use in ligating into pPDM. The primer of SEQ ID NO: 831 is an antisense primer with an added XhoI site for use in ligating into pPDM. The resulting fragment was cloned to pPDM at the NdeI and XhoI sites. Clones were confirmed by sequencing. For protein expression, the clone was

20 transformed into *E. coli* BL21 (DE3) CodonPlus-RIL competent cells. After induction, an OD600 of greater than 2.0 was achieved after 3 hours. Coomassie stained SDS-PAGE showed an over-expressed band at about 39 kD, and N-terminal sequencing confirmed the N-terminal to be that of P510S-E3. Optimized culture conditions are as follows: dilute overnight culture/daytime culture (LB + kanamycin + chloramphenicol) into 2x YT

25 (kanamycin and chloramphenicol) at a ratio of 25 ml culture to 1 liter 2x YT. Allow to grow at 37 °C until OD 600 equals 0.6. Take out an aliquot as T0 sample. Add 1 mM IPTG and allow to grow at 30 °C for 3 hours. Take out a T3 sample, spin down the cells and store at -80 °C until purification. The determined cDNA and amino acid sequences for the P510S-E3 construct are provided in SEQ ID NO: 824 and 827, respectively.

G) EXPRESSION OF P775S IN *E. COLI*

The antigen P775P contains multiple open reading frames (ORF). The third ORF, encoding the protein of SEQ ID NO: 483, has the best emotif score. An expression fusion construct containing the *M. tuberculosis* antigen Ra12 (SEQ ID NO: 819) and P775P-ORF3 with an N-terminal 6x HisTag was prepared as follows. P775P-ORF3 was amplified using the sense PCR primers of SEQ ID NO: 832 and the anti-sense PCR primer of SEQ ID NO: 833. The PCR amplified fragment of P775P and Ra12/pCRX1 were digested with the restriction enzymes EcoRI and XhoI. Vector and insert were ligated and then transformed into NovaBlue cells. Colonies were randomly screened for insert and then sequenced. A clone having the desired sequence was transformed into *E. coli* BL21 (DE3) CodonPlus-RIL competent cells. Two hours after induction, the cell density peaked at OD600 of approximately 1.8. Coomassie stained SDS-PAGE showed an over-expressed band at about 31 kD. Western blot using 6x HisTag antibody confirmed that the band was Ra12-P775P-ORF3. The determined cDNA and amino acid sequences for the fusion construct are provided in SEQ ID NO: 834 and 835, respectively.

EXAMPLE 18

PREPARATION AND CHARACTERIZATION OF ANTIBODIES

AGAINST PROSTATE-SPECIFIC POLYPEPTIDES

A) PREPARATION AND CHARACTERIZATION OF POLYCLONAL ANTIBODIES AGAINST P703P, P504S AND P509S

Polyclonal antibodies against P703P, P504S and P509S were prepared as follows.

Each prostate tumor antigen expressed in an *E. coli* recombinant expression system was grown overnight in LB broth with the appropriate antibiotics at 37°C in a shaking incubator. The next morning, 10 ml of the overnight culture was added to 500 ml

to 2x YT plus appropriate antibiotics in a 2L-baffled Erlenmeyer flask. When the Optical Density (at 560 nm) of the culture reached 0.4-0.6, the cells were induced with IPTG (1 mM). Four hours after induction with IPTG, the cells were harvested by centrifugation. The cells were then washed with phosphate buffered saline and centrifuged again. The supernatant was discarded and the cells were either frozen for future use or immediately processed. Twenty ml of lysis buffer was added to the cell pellets and vortexed. To break open the *E. coli* cells, this mixture was then run through the French Press at a pressure of 16,000 psi. The cells were then centrifuged again and the supernatant and pellet were checked by SDS-PAGE for the partitioning of the recombinant protein. For proteins that localized to the cell pellet, the pellet was resuspended in 10 mM Tris pH 8.0, 1% CHAPS and the inclusion body pellet was washed and centrifuged again. This procedure was repeated twice more. The washed inclusion body pellet was solubilized with either 8 M urea or 6 M guanidine HCl containing 10 mM Tris pH 8.0 plus 10 mM imidazole. The solubilized protein was added to 5 ml of nickel-chelate resin (Qiagen) and incubated for 45 min to 1 hour at room temperature with continuous agitation. After incubation, the resin and protein mixture were poured through a disposable column and the flow through was collected. The column was then washed with 10-20 column volumes of the solubilization buffer. The antigen was then eluted from the column using 8M urea, 10 mM Tris pH 8.0 and 300 mM imidazole and collected in 3 ml fractions. A SDS-PAGE gel was run to determine which fractions to pool for further purification.

As a final purification step, a strong anion exchange resin such as HiPrepQ (Biorad) was equilibrated with the appropriate buffer and the pooled fractions from above were loaded onto the column. Each antigen was eluted off the column with a increasing salt gradient. Fractions were collected as the column was run and another SDS-PAGE gel was run to determine which fractions from the column to pool. The pooled fractions were dialyzed against 10 mM Tris pH 8.0. The proteins were then vialled after filtration through a 0.22 micron filter and the antigens were frozen until needed for immunization.

Four hundred micrograms of each prostate antigen was combined with 100 micrograms of muramyl dipeptide (MDP). Every four weeks rabbits were boosted with 100

micrograms mixed with an equal volume of Incomplete Freund's Adjuvant (IFA). Seven days following each boost, the animal was bled. Sera was generated by incubating the blood at 4°C for 12-4 hours followed by centrifugation.

Ninety-six well plates were coated with antigen by incubating with 50 microliters (typically 1 microgram) of recombinant protein at 4 °C for 20 hours. 250 microliters of BSA blocking buffer was added to the wells and incubated at room temperature for 2 hours. Plates were washed 6 times with PBS/0.01% Tween. Rabbit sera was diluted in PBS. Fifty microliters of diluted sera was added to each well and incubated at room temperature for 30 min. Plates were washed as described above before 50 microliters of goat anti-rabbit horse radish peroxidase (HRP) at a 1:10000 dilution was added and incubated at room temperature for 30 min. Plates were again washed as described above and 100 microliters of TMB microwell peroxidase substrate was added to each well. Following a 15 min incubation in the dark at room temperature, the colorimetric reaction was stopped with 100 microliters of 1N H₂SO₄ and read immediately at 450 nm.

All polyclonal antibodies showed immunoreactivity to the appropriate antigen.

B) PREPARATION AND CHARACTERIZATION OF ANTIBODIES AGAINST P501S

A murine monoclonal antibody directed against the carboxy-terminus of the prostate-specific antigen P501S was prepared as follows.

A truncated fragment of P501S (amino acids 355-526 of SEQ ID NO: 113) was generated and cloned into the pET28b vector (Novagen) and expressed in *E. coli* as a thioredoxin fusion protein with a histidine tag. The trx-P501S fusion protein was purified by nickel chromatography, digested with thrombin to remove the trx fragment and further purified by an acid precipitation procedure followed by reverse phase HPLC.

Mice were immunized with truncated P501S protein. Serum bleeds from mice that potentially contained anti-P501S polyclonal sera were tested for P501S-specific reactivity using ELISA assays with purified P501S and trx-P501S proteins. Serum bleeds that appeared to react specifically with P501S were then screened for P501S reactivity by Western analysis. Mice that contained a P501S-specific antibody component were

sacrificed and spleen cells were used to generate anti-P501S antibody producing hybridomas using standard techniques. Hybridoma supernatants were tested for P501S-specific reactivity initially by ELISA, and subsequently by FACS analysis of reactivity with P501S transduced cells. Based on these results, a monoclonal hybridoma referred to as 10E3 was chosen for further subcloning. A number of subclones were generated, tested for specific reactivity to P501S using ELISA and typed for IgG isotype. The results of this analysis are shown below in Table V. Of the 16 subclones tested, the monoclonal antibody 10E3-G4-D3 was selected for further study.

10

Table V

Isotype analysis of murine anti-P501S monoclonal antibodies

Hybridoma clone	Isotype	Estimated [Ig] in supernatant (µg/ml)
4D11	IgG1	14.6
1G1	IgG1	0.6
4F6	IgG1	72
4H5	IgG1	13.8
4H5-E12	IgG1	10.7
4H5-EH2	IgG1	9.2
4H5-H2-A10	IgG1	10
4H5-H2-A3	IgG1	12.8
4H5-H2-A10-G6	IgG1	13.6
4H5-H2-B11	IgG1	12.3
10E3	IgG2a	3.4
10E3-D4	IgG2a	3.8
10E3-D4-G3	IgG2a	9.5
10E3-D4-G6	IgG2a	10.4
10E3-E7	IgG2a	6.5
8H12	IgG2a	0.6

The specificity of 10E3-G4-D3 for P501S was examined by FACS analysis. Specifically, cells were fixed (2% formaldehyde, 10 minutes), permeabilized (0.1% saponin, 10 minutes) and stained with 10E3-G4-D3 at 0.5 – 1 µg/ml, followed by incubation with a secondary, FITC-conjugated goat anti-mouse Ig antibody (Pharmingen, San Diego, CA). Cells were then analyzed for FITC fluorescence using an Excalibur

fluorescence activated cell sorter. For FACS analysis of transduced cells, B-LCL were retrovirally transduced with P501S. For analysis of infected cells, B-LCL were infected with a vaccinia vector that expresses P501S. To demonstrate specificity in these assays, B-LCL transduced with a different antigen (P703P) and uninfected B-LCL vectors were
 5 utilized. 10E3-G4-D3 was shown to bind with P501S-transduced B-LCL and also with P501S-infected B-LCL, but not with either uninfected cells or P703P-transduced cells.

To determine whether the epitope recognized by 10E3-G4-D3 was found on the surface or in an intracellular compartment of cells, B-LCL were transduced with P501S or HLA-B8 as a control antigen and either fixed and permeabilized as described above or
 10 directly stained with 10E3-G4-D3 and analyzed as above. Specific recognition of P501S by 10E3-G4-D3 was found to require permeabilization, suggesting that the epitope recognized by this antibody is intracellular.

The reactivity of 10E3-G4-D3 with the three prostate tumor cell lines Lncap, PC-3 and DU-145, which are known to express high, medium and very low levels
 15 of P501S, respectively, was examined by permeabilizing the cells and treating them as described above. Higher reactivity of 10E3-G4-D3 was seen with Lncap than with PC-3, which in turn showed higher reactivity than DU-145. These results are in agreement with the real time PCR and demonstrate that the antibody specifically recognizes P501S in these tumor cell lines and that the epitope recognized in prostate tumor cell lines is also
 20 intracellular.

Specificity of 10E3-G4-D3 for P501S was also demonstrated by Western blot analysis. Lysates from the prostate tumor cell lines Lncap, DU-145 and PC-3, from P501S-transiently transfected HEK293 cells, and from non-transfected HEK293 cells were generated. Western blot analysis of these lysates with 10E3-G4-D3 revealed a 46 kDa
 25 immunoreactive band in Lncap, PC-3 and P501S-transfected HEK cells, but not in DU-145 cells or non-transfected HEK293 cells. P501S mRNA expression is consistent with these results since semi-quantitative PCR analysis revealed that P501S mRNA is expressed in Lncap, to a lesser but detectable level in PC-3 and not at all in DU-145 cells. Bacterially expressed and purified recombinant P501S (referred to as P501SStr2) was recognized by

10E3-G4-D3 (24 kDa), as was full-length P501S that was transiently expressed in HEK293 cells using either the expression vector VR1012 or pCEP4. Although the predicted molecular weight of P501S is 60.5 kDa, both transfected and “native” P501S run at a slightly lower mobility due to its hydrophobic nature.

5 Immunohistochemical analysis was performed on prostate tumor and a panel of normal tissue sections (prostate, adrenal, breast, cervix, colon, duodenum, gall bladder, ileum, kidney, ovary, pancreas, parotid gland, skeletal muscle, spleen and testis). Tissue samples were fixed in formalin solution for 24 hours and embedded in paraffin before being sliced into 10 micron sections. Tissue sections were permeabilized and
10 incubated with 10E3-G4-D3 antibody for 1 hr. HRP-labeled anti-mouse followed by incubation with DAB chromogen was used to visualize P501S immunoreactivity. P501S was found to be highly expressed in both normal prostate and prostate tumor tissue but was not detected in any of the other tissues tested.

To identify the epitope recognized by 10E3-G4-D3, an epitope mapping
15 approach was pursued. A series of 13 overlapping 20-21 mers (5 amino acid overlap; SEQ ID NO: 489-501) was synthesized that spanned the fragment of P501S used to generate 10E3-G4-D3. Flat bottom 96 well microtiter plates were coated with either the peptides or the P501S fragment used to immunize mice, at 1 microgram/ml for 2 hours at 37 °C. Wells were then aspirated and blocked with phosphate buffered saline containing 1% (w/v) BSA
20 for 2 hours at room temperature, and subsequently washed in PBS containing 0.1% Tween 20 (PBST). Purified antibody 10E3-G4-D3 was added at 2 fold dilutions (1000 ng – 16 ng) in PBST and incubated for 30 minutes at room temperature. This was followed by washing 6 times with PBST and subsequently incubating with HRP-conjugated donkey anti-mouse IgG (H+L)Affinipure F(ab') fragment (Jackson ImmunoResearch, West Grove, PA) at
25 1:20000 for 30 minutes. Plates were then washed and incubated for 15 minutes in tetramethyl benzidine. Reactions were stopped by the addition of 1N sulfuric acid and plates were read at 450 nm using an ELISA plate reader. As shown in Fig. 8, reactivity was seen with the peptide of SEQ ID NO: 496 (corresponding to amino acids 439-459 of P501S) and with the P501S fragment but not with the remaining peptides, demonstrating

that the epitope recognized by 10E3-G4-D3 is localized to amino acids 439-459 of SEQ ID NO: 113.

In order to further evaluate the tissue specificity of P501S, multi-array immunohistochemical analysis was performed on approximately 4700 different human tissues encompassing all the major normal organs as well as neoplasias derived from these tissues. Sixty-five of these human tissue samples were of prostate origin. Tissue sections 0.6 mm in diameter were formalin-fixed and paraffin embedded. Samples were pretreated with HIER using 10 mM citrate buffer pH 6.0 and boiling for 10 min. Sections were stained with 10E3-G4-D3 and P501S immunoreactivity was visualized with HRP. All the 65 prostate tissues samples (5 normal, 55 untreated prostate tumors, 5 hormone refractory prostate tumors) were positive, showing distinct perinuclear staining. All other tissues examined were negative for P501S expression.

C) PREPARATION AND CHARACTERIZATION OF ANTIBODIES AGAINST P503S

A fragment of P503S (amino acids 113-241 of SEQ ID NO: 114) was expressed and purified from bacteria essentially as described above for P501S and used to immunize both rabbits and mice. Mouse monoclonal antibodies were isolated using standard hybridoma technology as described above. Rabbit monoclonal antibodies were isolated using Selected Lymphocyte Antibody Method (SLAM) technology at Immgenics Pharmaceuticals (Vancouver, BC, Canada). Table VI, below, lists the monoclonal antibodies that were developed against P503S.

Table VI

Antibody	Species
20D4	Rabbit
JA1	Rabbit
1A4	Mouse
1C3	Mouse
1C9	Mouse
1D12	Mouse

Antibody	Species
2A11	Mouse
2H9	Mouse
4H7	Mouse
8A8	Mouse
8D10	Mouse
9C12	Mouse
6D12	Mouse

The DNA sequences encoding the complementarity determining regions (CDRs) for the rabbit monoclonal antibodies 20D4 and JA1 were determined and are provided in SEQ ID NO: 502 and 503, respectively.

5 In order to better define the epitope binding region of each of the antibodies, a series of overlapping peptides were generated that span amino acids 109-213 of SEQ ID NO: 114. These peptides were used to epitope map the anti-P503S monoclonal antibodies by ELISA as follows. The recombinant fragment of P503S that was employed as the immunogen was used as a positive control. Ninety-six well microtiter plates were coated
10 with either peptide or recombinant antigen at 20 ng/well overnight at 4 °C. Plates were aspirated and blocked with phosphate buffered saline containing 1% (w/v) BSA for 2 hours at room temperature then washed in PBS containing 0.1% Tween 20 (PBST). Purified rabbit monoclonal antibodies diluted in PBST were added to the wells and incubated for 30 min at room temperature. This was followed by washing 6 times with PBST and
15 incubation with Protein-A HRP conjugate at a 1:2000 dilution for a further 30 min. Plates were washed six times in PBST and incubated with tetramethylbenzidine (TMB) substrate for a further 15 min. The reaction was stopped by the addition of 1N sulfuric acid and plates were read at 450 nm using at ELISA plate reader. ELISA with the mouse monoclonal antibodies was performed with supernatants from tissue culture run neat in the
20 assay.

All of the antibodies bound to the recombinant P503S fragment, with the exception of the negative control SP2 supernatant. 20D4, JA1 and 1D12 bound strictly to peptide #2101 (SEQ ID NO: 504), which corresponds to amino acids 151-169 of SEQ ID

NO: 114. 1C3 bound to peptide #2102 (SEQ ID NO: 505), which corresponds to amino acids 165-184 of SEQ ID NO: 114. 9C12 bound to peptide #2099 (SEQ ID NO: 522), which corresponds to amino acids 120-139 of SEQ ID NO: 114. The other antibodies bind to regions that were not examined in these studies.

5 Subsequent to epitope mapping, the antibodies were tested by FACS analysis on a cell line that stably expressed P503S to confirm that the antibodies bind to cell surface epitopes. Cells stably transfected with a control plasmid were employed as a negative control. Cells were stained live with no fixative. 0.5 ug of anti-P503S monoclonal antibody was added and cells were incubated on ice for 30 min before being
10 washed twice and incubated with a FITC-labelled goat anti-rabbit or mouse secondary antibody for 20 min. After being washed twice, cells were analyzed with an Excalibur fluorescent activated cell sorter. The monoclonal antibodies 1C3, 1D12, 9C12, 20D4 and JA1, but not 8D3, were found to bind to a cell surface epitope of P503S.

In order to determine which tissues express P503S, immunohistochemical
15 analysis was performed, essentially as described above, on a panel of normal tissues (prostate, adrenal, breast, cervix, colon, duodenum, gall bladder, ileum, kidney, ovary, pancreas, parotid gland, skeletal muscle, spleen and testis). HRP-labeled anti-mouse or anti-rabbit antibody followed by incubation with TMB was used to visualize P503S immunoreactivity. P503S was found to be highly expressed in prostate tissue, with lower
20 levels of expression being observed in cervix, colon, ileum and kidney, and no expression being observed in adrenal, breast, duodenum, gall bladder, ovary, pancreas, parotid gland, skeletal muscle, spleen and testis.

Western blot analysis was used to characterize anti-P503S monoclonal antibody specificity. SDS-PAGE was performed on recombinant (rec) P503S expressed in
25 and purified from bacteria and on lysates from HEK293 cells transfected with full length P503S. Protein was transferred to nitrocellulose and then Western blotted with each of the anti-P503S monoclonal antibodies (20D4, JA1, 1D12, 6D12 and 9C12) at an antibody concentration of 1 ug/ml. Protein was detected using horse radish peroxidase (HRP) conjugated to either a goat anti-mouse monoclonal antibody or to protein A-sepharose.

The monoclonal antibody 20D4 detected the appropriate molecular weight 14 kDa recombinant P503S (amino acids 113-241) and the 23.5 kDa species in the HEK293 cell lysates transfected with full length P503S. Other anti-P503S monoclonal antibodies displayed similar specificity by Western blot.

5 **D) PREPARATION AND CHARACTERIZATION OF ANTIBODIES AGAINST P703P**

Rabbits were immunized with either a truncated (P703Ptr1; SEQ ID NO: 172) or full-length mature form (P703Pfl; SEQ ID NO: 523) of recombinant P703P protein was expressed in and purified from bacteria as described above. Affinity purified polyclonal antibody was generated using immunogen P703Pfl or P703Ptr1 attached to a solid support. Rabbit monoclonal antibodies were isolated using SLAM technology at Immgenics Pharmaceuticals. Table VII below lists both the polyclonal and monoclonal antibodies that were generated against P703P.

Table VII

Antibody	Immunogen	Species/type
Aff. Purif. P703P (truncated); #2594	P703Ptrl	Rabbit polyclonal
Aff. Purif. P703P (full length); #9245	P703Pfl	Rabbit polyclonal
2D4	P703Ptrl	Rabbit monoclonal
8H2	P703Ptrl	Rabbit monoclonal
7H8	P703Ptrl	Rabbit monoclonal

The DNA sequences encoding the complementarity determining regions (CDRs) for the rabbit monoclonal antibodies 8H2, 7H8 and 2D4 were determined and are provided in SEQ ID NO: 506-508, respectively.

Epitope mapping studies were performed as described above. Monoclonal antibodies 2D4 and 7H8 were found to specifically bind to the peptides of SEQ ID NO: 509 (corresponding to amino acids 145-159 of SEQ ID NO: 172) and SEQ ID NO: 510 (corresponding to amino acids 11-25 of SEQ ID NO: 172), respectively. The polyclonal

antibody 2594 was found to bind to the peptides of SEQ ID NO: 511-514, with the polyclonal antibody 9427 binding to the peptides of SEQ ID NO: 515-517.

The specificity of the anti-P703P antibodies was determined by Western blot analysis as follows. SDS-PAGE was performed on (1) bacterially expressed recombinant antigen; (2) lysates of HEK293 cells and Ltk^{-/-} cells either untransfected or transfected with a plasmid expressing full length P703P; and (3) supernatant isolated from these cell cultures. Protein was transferred to nitrocellulose and then Western blotted using the anti-P703P polyclonal antibody #2594 at an antibody concentration of 1 ug/ml. Protein was detected using horse radish peroxidase (HRP) conjugated to an anti-rabbit antibody. A 35 kDa immunoreactive band could be observed with recombinant P703P. Recombinant P703P runs at a slightly higher molecular weight since it is epitope tagged. In lysates and supernatants from cells transfected with full length P703P, a 30 kDa band corresponding to P703P was observed. To assure specificity, lysates from HEK293 cells stably transfected with a control plasmid were also tested and were negative for P703P expression. Other anti-P703P antibodies showed similar results.

Immunohistochemical studies were performed as described above, using anti-P703P monoclonal antibody. P703P was found to be expressed at high levels in normal prostate and prostate tumor tissue but was not detectable in all other tissues tested (breast tumor, lung tumor and normal kidney).

EXAMPLE 19

CHARACTERIZATION OF CELL SURFACE EXPRESSION AND CHROMOSOME LOCALIZATION OF THE PROSTATE-SPECIFIC ANTIGEN P501S

This example describes studies demonstrating that the prostate-specific antigen P501S is expressed on the surface of cells, together with studies to determine the probable chromosomal location of P501S.

The protein P501S (SEQ ID NO: 113) is predicted to have 11 transmembrane domains. Based on the discovery that the epitope recognized by the anti-

P501S monoclonal antibody 10E3-G4-D3 (described above in Example 17) is intracellular, it was predicted that following transmembrane determinants would allow the prediction of extracellular domains of P501S. Fig. 9 is a schematic representation of the P501S protein showing the predicted location of the transmembrane domains and the intracellular epitope described in Example 17. Underlined sequence represents the predicted transmembrane domains, bold sequence represents the predicted extracellular domains, and italicized sequence represents the predicted intracellular domains. Sequence that is both bold and underlined represents sequence employed to generate polyclonal rabbit serum. The location of the transmembrane domains was predicted using HHMTOP as described by Tusnady and Simon (Principles Governing Amino Acid Composition of Integral Membrane Proteins: Applications to Topology Prediction, *J. Mol. Biol.* 283:489-506, 1998).

Based on Fig. 9, the P501S domain flanked by the transmembrane domains corresponding to amino acids 274-295 and 323-342 is predicted to be extracellular. The peptide of SEQ ID NO: 518 corresponds to amino acids 306-320 of P501S and lies in the predicted extracellular domain. The peptide of SEQ ID NO: 519, which is identical to the peptide of SEQ ID NO: 518 with the exception of the substitution of the histidine with an asparagine, was synthesized as described above. A Cys-Gly was added to the C-terminus of the peptide to facilitate conjugation to the carrier protein. Cleavage of the peptide from the solid support was carried out using the following cleavage mixture: trifluoroacetic acid:ethanediol:thioanisole:water:phenol (40:1:2:2:3). After cleaving for two hours, the peptide was precipitated in cold ether. The peptide pellet was then dissolved in 10% v/v acetic acid and lyophilized prior to purification by C18 reverse phase hplc. A gradient of 5-60% acetonitrile (containing 0.05% TFA) in water (containing 0.05% TFA) was used to elute the peptide. The purity of the peptide was verified by hplc and mass spectrometry, and was determined to be >95%. The purified peptide was used to generate rabbit polyclonal antisera as described above.

Surface expression of P501S was examined by FACS analysis. Cells were stained with the polyclonal anti-P501S peptide serum at 10 µg/ml, washed, incubated with

a secondary FITC-conjugated goat anti-rabbit Ig antibody (ICN), washed and analyzed for FITC fluorescence using an Excalibur fluorescence activated cell sorter. For FACS analysis of transduced cells, B-LCL were retrovirally transduced with P501S. To demonstrate specificity in these assays, B-LCL transduced with an irrelevant antigen (P703P) or nontransduced were stained in parallel. For FACS analysis of prostate tumor cell lines, Lncap, PC-3 and DU-145 were utilized. Prostate tumor cell lines were dissociated from tissue culture plates using cell dissociation medium and stained as above. All samples were treated with propidium iodide (PI) prior to FACS analysis, and data was obtained from PI-excluding (*i.e.*, intact and non-permeabilized) cells. The rabbit polyclonal serum generated against the peptide of SEQ ID NO: 519 was shown to specifically recognize the surface of cells transduced to express P501S, demonstrating that the epitope recognized by the polyclonal serum is extracellular.

To determine biochemically if P501S is expressed on the cell surface, peripheral membranes from Lncap cells were isolated and subjected to Western blot analysis. Specifically, Lncap cells were lysed using a dounce homogenizer in 5 ml of homogenization buffer (250 mM sucrose, 10 mM HEPES, 1mM EDTA, pH 8.0, 1 complete protease inhibitor tablet (Boehringer Mannheim)). Lysate samples were spun at 1000 g for 5 min at 4 °C. The supernatant was then spun at 8000g for 10 min at 4 °C. Supernatant from the 8000g spin was recovered and subjected to a 100,000g spin for 30 min at 4 °C to recover peripheral membrane. Samples were then separated by SDS-PAGE and Western blotted with the mouse monoclonal antibody 10E3-G4-D3 (described above in Example 17) using conditions described above. Recombinant purified P501S, as well as HEK293 cells transfected with and over-expressing P501S were included as positive controls for P501S detection. LCL cell lysate was included as a negative control. P501S could be detected in Lncap total cell lysate, the 8000g (internal membrane) fraction and also in the 100,000g (plasma membrane) fraction. These results indicate that P501S is expressed at, and localizes to, the peripheral membrane.

To demonstrate that the rabbit polyclonal antiserum generated to the peptide of SEQ ID NO: 519 specifically recognizes this peptide as well as the corresponding native

peptide of SEQ ID NO: 518, ELISA analyses were performed. For these analyses, flat-bottomed 96 well microtiter plates were coated with either the peptide of SEQ ID NO: 519, the longer peptide of SEQ ID NO: 520 that spans the entire predicted extracellular domain, the peptide of SEQ ID NO: 521 which represents the epitope recognized by the P501S-specific antibody 10E3-G4-D3, or a P501S fragment (corresponding to amino acids 355-526 of SEQ ID NO: 113) that does not include the immunizing peptide sequence, at 1 µg/ml for 2 hours at 37 °C. Wells were aspirated, blocked with phosphate buffered saline containing 1% (w/v) BSA for 2 hours at room temperature and subsequently washed in PBS containing 0.1% Tween 20 (PBST). Purified anti-P501S polyclonal rabbit serum was added at 2 fold dilutions (1000 ng - 125 ng) in PBST and incubated for 30 min at room temperature. This was followed by washing 6 times with PBST and incubating with HRP-conjugated goat anti-rabbit IgG (H+L) Affinipure F(ab') fragment at 1:20000 for 30 min. Plates were then washed and incubated for 15 min in tetramethyl benzidine. Reactions were stopped by the addition of 1N sulfuric acid and plates were read at 450 nm using an ELISA plate reader. As shown in Fig. 11, the anti-P501S polyclonal rabbit serum specifically recognized the peptide of SEQ ID NO: 519 used in the immunization as well as the longer peptide of SEQ ID NO: 520, but did not recognize the irrelevant P501S-derived peptides and fragments.

In further studies, rabbits were immunized with peptides derived from the P501S sequence and predicted to be either extracellular or intracellular, as shown in Fig. 9. Polyclonal rabbit sera were isolated and polyclonal antibodies in the serum were purified, as described above. To determine specific reactivity with P501S, FACS analysis was employed, utilizing either B-LCL transduced with P501S or the irrelevant antigen P703P, of B-LCL infected with vaccinia virus-expressing P501S. For surface expression, dead and non-intact cells were excluded from the analysis as described above. For intracellular staining, cells were fixed and permeabilized as described above. Rabbit polyclonal serum generated against the peptide of SEQ ID NO: 548, which corresponds to amino acids 181-198 of P501S, was found to recognize a surface epitope of P501S. Rabbit polyclonal serum generated against the peptide SEQ ID NO: 551, which corresponds to amino acids

543-553 of P501S, was found to recognize an epitope that was either potentially extracellular or intracellular since in different experiments intact or permeabilized cells were recognized by the polyclonal sera. Based on similar deductive reasoning, the sequences of SEQ ID NO: 541-547, 549 and 550, which correspond to amino acids 109-122, 539-553, 509-520, 37-54, 342-359, 295-323, 217-274, 143-160 and 75-88, respectively, of P501S, can be considered to be potential surface epitopes of P501S recognized by antibodies.

The chromosomal location of P501S was determined using the GeneBridge 4 Radiation Hybrid panel (Research Genetics). The PCR primers of SEQ ID NO: 528 and 529 were employed in PCR with DNA pools from the hybrid panel according to the manufacturer's directions. After 38 cycles of amplification, the reaction products were separated on a 1.2% agarose gel, and the results were analyzed through the Whitehead Institute/MIT Center for Genome Research web server (<http://www-genome.wi.mit.edu/cgi-bin/contig/rhmapper.pl>) to determine the probable chromosomal location. Using this approach, P501S was mapped to the long arm of chromosome 1 at WI-9641 between q32 and q42. This region of chromosome 1 has been linked to prostate cancer susceptibility in hereditary prostate cancer (Smith *et al. Science* 274:1371-1374, 1996 and Berthon *et al. Am. J. Hum. Genet.* 62:1416-1424, 1998). These results suggest that P501S may play a role in prostate cancer malignancy.

20

EXAMPLE 20

REGULATION OF EXPRESSION OF THE PROSTATE-SPECIFIC ANTIGEN P501S

Steroid (androgen) hormone modulation is a common treatment modality in prostate cancer. The expression of a number of prostate tissue-specific antigens have previously been demonstrated to respond to androgen. The responsiveness of the prostate-specific antigen P501S to androgen treatment was examined in a tissue culture system as follows.

Cells from the prostate tumor cell line LNCaP were plated at 1.5×10^6 cells/T75 flask (for RNA isolation) or 3×10^5 cells/well of a 6-well plate (for FACS analysis) and grown overnight in RPMI 1640 media containing 10% charcoal-stripped fetal calf serum (BRL Life Technologies, Gaithersburg, MD). Cell culture was continued for an additional 72 hours in RPMI 1640 media containing 10% charcoal-stripped fetal calf serum, with 1 nM of the synthetic androgen Methyltrienolone (R1881; New England Nuclear) added at various time points. Cells were then harvested for RNA isolation and FACS analysis at 0, 1, 2, 4, 8, 16, 24, 28 and 72-hours post androgen addition. FACS analysis was performed using the anti-P501S antibody 10E3-G4-D3 and permeabilized cells.

For Northern analysis, 5-10 micrograms of total RNA was run on a formaldehyde denaturing gel, transferred to Hybond-N nylon membrane (Amersham Pharmacia Biotech, Piscataway, NJ), cross-linked and stained with methylene blue. The filter was then prehybridized with Church's Buffer (250 mM Na_2HPO_4 , 70 mM H_3PO_4 , 1 mM EDTA, 1% SDS, 1% BSA in pH 7.2) at 65 °C for 1 hour. P501S DNA was labeled with ^{32}P using High Prime random-primed DNA labeling kit (Boehringer Mannheim). Unincorporated label was removed using MicroSpin S300-HR columns (Amersham Pharmacia Biotech). The RNA filter was then hybridized with fresh Church's Buffer containing labeled cDNA overnight, washed with 1X SCP (0.1 M NaCl, 0.03 M $\text{Na}_2\text{HPO}_4 \cdot 7\text{H}_2\text{O}$, 0.001 M Na_2EDTA), 1% sarkosyl (n-lauroylsarcosine) and exposed to X-ray film.

Using both FACS and Northern analysis, P501S message and protein levels were found to increase in response to androgen treatment.

25

EXAMPLE 20

PREPARATION OF FUSION PROTEINS OF PROSTATE-SPECIFIC ANTIGENS

The example describes the preparation of a fusion protein of the prostate-specific antigen P703P and a truncated form of the known prostate antigen PSA. The

truncated form of PSA has a 21 amino acid deletion around the active serine site. The expression construct for the fusion protein also has a restriction site at 3' end, immediately prior to the termination codon, to aid in adding cDNA for additional antigens.

The full-length cDNA for PSA was obtained by RT-PCR from a pool of
 5 RNA from human prostate tumor tissues using the primers of SEQ ID NO: 607 and 608, and cloned in the vector pCR-Blunt II-TOPO. The resulting cDNA was employed as a template to make two different fragments of PSA by PCR with two sets of primers (SEQ ID NO: 609 and 610; and SEQ ID NO: 611 and 612). The PCR products having the expected size were used as templates to make truncated forms of PSA by PCR with the
 10 primers of SEQ ID NO: 611 and 613, which generated PSA (delta 208-218 in amino acids). The cDNA for the mature form of P703P with a 6X histidine tag at the 5' end, was prepared by PCR with P703P and the primers of SEQ ID NO: 614 and 615. The cDNA for the fusion of P703P with the truncated form of PSA (referred to as FOPP) was then obtained by PCR using the modified P703P cDNA and the truncated form of PSA cDNA
 15 as templates and the primers of SEQ ID NO: 614 and 615. The FOPP cDNA was cloned into the NdeI site and XhoI site of the expression vector pCRX1, and confirmed by DNA sequencing. The determined cDNA sequence for the fusion construct FOPP is provided in SEQ ID NO: 616, with the amino acid sequence being provided in SEQ ID NO: 617.

The fusion FOPP was expressed as a single recombinant protein in *E. coli* as
 20 follows. The expression plasmid pCRX1FOPP was transformed into the *E. coli* strain BL21-CodonPlus RIL. The transformant was shown to express FOPP protein upon induction with 1 mM IPTG. The culture of the corresponding expression clone was inoculated into 25 ml LB broth containing 50 ug/ml kanamycin and 34 ug/ml chloramphenicol, grown at 37 °C to OD600 of about 1, and stored at 4 °C overnight. The
 25 culture was diluted into 1 liter of TB LB containing 50 ug/ml kanamycin and 34 ug/ml chloramphenicol, and grown at 37 °C to OD600 of 0.4. IPTG was added to a final concentration of 1 mM, and the culture was incubated at 30 °C for 3 hours. The cells were pelleted by centrifugation at 5,000 RPM for 8 min. To purify the protein, the cell pellet was suspended in 25 ml of 10 mM Tris-Cl pH 8.0, 2mM PMSF, complete protease

inhibitor and 15 ug lysozyme. The cells were lysed at 4 °C for 30 minutes, sonicated several times and the lysate centrifuged for 30 minutes at 10,000 x g. The precipitate, which contained the inclusion body, was washed twice with 10 mM Tris-Cl pH 8.0 and 1% CHAPS. The inclusion body was dissolved in 40 ml of 10 mM Tris-Cl pH 8.0, 100 mM sodium phosphate and 8 M urea. The solution was bound to 8 ml Ni-NTA (Qiagen) for one hour at room temperature. The mixture was poured into a 25 ml column and washed with 50 ml of 10 mM Tris-Cl pH 6.3, 100 mM sodium phosphate, 0.5% DOC and 8M urea. The bound protein was eluted with 350 mM imidazole, 10 mM Tris-Cl pH 8.0, 100 mM sodium phosphate and 8 M urea. The fractions containing FOPP proteins were combined and dialyzed extensively against 10 mM Tris-Cl pH 4.6, aliquoted and stored at - 70 °C.

EXAMPLE 21

REAL-TIME PCR CHARACTERIZATION OF THE PROSTATE-SPECIFIC ANTIGEN P501S IN PERIPHERAL BLOOD OF PROSTATE CANCER PATIENTS

Circulating epithelial cells were isolated from fresh blood of normal individuals and metastatic prostate cancer patients, mRNA isolated and cDNA prepared using real-time PCR procedures. Real-time PCR was performed with the TaqmanTM procedure using both gene specific primers and probes to determine the levels of gene expression.

Epithelial cells were enriched from blood samples using an immunomagnetic bead separation method (Dynal A.S., Oslo, Norway). Isolated cells were lysed and the magnetic beads removed. The lysate was then processed for poly A+ mRNA isolation using magnetic beads coated with Oligo(dT)25. After washing the beads in buffer, bead/poly A+ RNA samples were suspended in 10 mM Tris HCl pH 8.0 and subjected to reversed transcription. The resulting cDNA was subjected to real-time PCR using gene specific primers. Beta-actin content was also determined and used for normalization. Samples with P501S copies greater than the mean of the normal samples + 3 standard deviations were considered positive. Real time PCR on blood samples was

CLAIMS

What is claimed:

1. An isolated polypeptide, comprising at least an immunogenic portion of a prostate-specific protein, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(a) sequences recited in SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772, 779, 817, 823 and 824;

(b) sequences that hybridize to a sequence recited in any one of SEQ ID NOs: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772, 779, 817, 823 and 824 under moderately stringent conditions; and

(c) complements of sequences of (a) or (b).

2. An isolated polypeptide according to claim 1, wherein the polypeptide comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434,

435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772, 779, 817, 823 and 824, or a complement of any of the foregoing polynucleotide sequences.

3. An isolated polypeptide comprising a sequence recited in any one of SEQ ID NO: 108, 112, 113, 114, 172, 176, 178, 327, 329, 331, 338, 339, 383, 477-483, 496, 504, 505, 519, 520, 522, 525, 527, 532, 534, 537-551, 553-568, 573-586, 588-590, 592, 706-708, 778, 780, 781, 810, 811, 814, 818, 826 and 827.

4. An isolated polynucleotide encoding at least 15 amino acid residues of a prostate-specific protein, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide comprising a sequence recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772, 779, 817, 823 and 824, or a complement of any of the foregoing sequences.

5. An isolated polynucleotide encoding a prostate-specific protein, or a variant thereof, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide comprising a sequence recited in any one of SEQ ID NOs: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-

461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772, 779, 817, 823 and 824, or a complement of any of the foregoing sequences.

6. An isolated polynucleotide, comprising a sequence recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772, 779, 817, 823 and 824.

7. An isolated polynucleotide, comprising a sequence that hybridizes to a sequence recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772, 779, 817, 823 and 824 under moderately stringent conditions.

8. An isolated polynucleotide complementary to a polynucleotide according to any one of claims 4-7.

9. An expression vector, comprising a polynucleotide according to any one of claims 4-8.

10. A host cell transformed or transfected with an expression vector according to claim 9.

11. An isolated antibody, or antigen-binding fragment thereof, that specifically binds to a prostate-specific protein that comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772, 779, 817, 823 and 824, or a complement of any of the foregoing polynucleotide sequences.

12. A fusion protein, comprising at least one polypeptide according to claim 1.

13. A fusion protein according to claim 12, wherein the fusion protein comprises an expression enhancer that increases expression of the fusion protein in a host cell transfected with a polynucleotide encoding the fusion protein.

14. A fusion protein according to claim 12, wherein the fusion protein comprises a T helper epitope that is not present within the polypeptide of claim 1.

15. A fusion protein according to claim 12, wherein the fusion protein comprises an affinity tag.

16. An isolated polynucleotide encoding a fusion protein according to claim 12.

17. A pharmaceutical composition, comprising a physiologically acceptable carrier and at least one component selected from the group consisting of:

(a) a polypeptide according to claim 1;

- (b) a polynucleotide according to claim 4;
- (c) an antibody according to claim 11;
- (d) a fusion protein according to claim 12; and
- (e) a polynucleotide according to claim 16.

18. An immunogenic composition comprising an immunostimulant and at least one component selected from the group consisting of:

- (a) a polypeptide according to claim 1;
- (b) a polynucleotide according to claim 4;
- (c) an antibody according to claim 11;
- (d) a fusion protein according to claim 12; and
- (e) a polynucleotide according to claim 16.

19. An immunogenic composition according to claim 18, wherein the immunostimulant is an adjuvant.

20. An immunogenic composition according to claim 18, wherein the immunostimulant induces a predominantly Type I response.

21. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a pharmaceutical composition according to claim 17.

22. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of an immunogenic composition according to claim 18.

23. A pharmaceutical composition comprising an antigen-presenting cell that expresses a polypeptide according to claim 1, in combination with a pharmaceutically acceptable carrier or excipient.

24. A pharmaceutical composition according to claim 23, wherein the antigen presenting cell is a dendritic cell or a macrophage.

25. An immunogenic composition comprising an antigen-presenting cell that expresses a polypeptide comprising at least an immunogenic portion of a prostate-specific protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(a) sequences recited in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824;

(b) sequences that hybridize to a sequence recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824 under moderately stringent conditions; and

(c) complements of sequences of (i) or (ii);
in combination with an immunostimulant.

26. An immunogenic composition according to claim 25, wherein the immunostimulant is an adjuvant.

27. An immunogenic composition according to claim 25, wherein the immunostimulant induces a predominantly Type I response.

28. An immunogenic composition according to claim 25, wherein the antigen-presenting cell is a dendritic cell.

29. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of an antigen-presenting cell that expresses a polypeptide comprising at least an immunogenic portion of a prostate-specific protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(a) sequences recited in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824;

(b) sequences that hybridize to a sequence recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824 under moderately stringent conditions; and

(c) complements of sequences of (i) or (ii) encoded by a polynucleotide recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824;

and thereby inhibiting the development of a cancer in the patient.

30. A method according to claim 29, wherein the antigen-presenting cell is a dendritic cell.

31. A method according to any one of claims 21, 22 and 29, wherein the cancer is prostate cancer.

32. A method for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a prostate-specific protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824; and

(ii) complements of the foregoing polynucleotides;

wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the antigen from the sample.

33. A method according to claim 32, wherein the biological sample is blood or a fraction thereof.

34. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated according to the method of claim 32.

35. A method for stimulating and/or expanding T cells specific for a prostate-specific protein, comprising contacting T cells with at least one component selected from the group consisting of:

(a) polypeptides comprising at least an immunogenic portion of a prostate-specific protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) sequences recited in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824;

(ii) sequences that hybridize to a sequence recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824 under moderately stringent conditions; and

(iii) complements of sequences of (i) or (ii);

- (b) polynucleotides encoding a polypeptide of (a); and
- (c) antigen presenting cells that express a polypeptide of (a);

under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells.

36. An isolated T cell population, comprising T cells prepared according to the method of claim 35.

37. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a T cell population according to claim 36.

38. A method for inhibiting the development of a cancer in a patient, comprising the steps of:

(a) incubating CD4⁺ and/or CD8⁺ T cells isolated from a patient with at least one component selected from the group consisting of:

(i) polypeptides comprising at least an immunogenic portion of a prostate-specific protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(1) sequences recited in SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824;

(2) sequences that hybridize to a sequence recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824 under moderately stringent conditions; and

(3) complements of sequences of (1) or (2);

(ii) polynucleotides encoding a polypeptide of (i); and

(iii) antigen presenting cells that expresses a polypeptide of (i);
such that T cells proliferate; and

(b) administering to the patient an effective amount of the proliferated T cells,
and thereby inhibiting the development of a cancer in the patient.

39. A method for inhibiting the development of a cancer in a patient,
comprising the steps of:

(a) incubating CD4⁺ and/or CD8⁺ T cells isolated from a patient with at least
one component selected from the group consisting of:

(i) polypeptides comprising at least an immunogenic portion of a
prostate-specific protein, or a variant thereof, wherein the tumor protein comprises an amino acid
sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(1) sequences recited in SEQ ID NO: 1-111, 115-171, 173-175,
177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530,
531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and
824;

(2) sequences that hybridize to a sequence recited in any one of
SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375,
381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-
705, 709-774, 777, 789, 817, 823 and 824 under moderately stringent conditions; and

(3) complements of sequences of (1) or (2);

(ii) polynucleotides encoding a polypeptide of (i); and

(iii) antigen presenting cells that express a polypeptide of (i);

such that T cells proliferate;

(b) cloning at least one proliferated cell to provide cloned T cells; and

(c) administering to the patient an effective amount of the cloned T cells, and
thereby inhibiting the development of a cancer in the patient.

40. A method for determining the presence or absence of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with a binding agent that binds to a prostate-specific protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824 or a complement of any of the foregoing polynucleotide sequences;

(b) detecting in the sample an amount of polypeptide that binds to the binding agent; and

(c) comparing the amount of polypeptide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

41. A method according to claim 40, wherein the binding agent is an antibody.

42. A method according to claim 43, wherein the antibody is a monoclonal antibody.

43. A method according to claim 40, wherein the cancer is prostate cancer.

44. A method for monitoring the progression of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a prostate-specific protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824 or a complement of any of the foregoing polynucleotide sequences;

- (b) detecting in the sample an amount of polypeptide that binds to the binding agent;
- (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and
- (d) comparing the amount of polypeptide detected in step (c) to the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

45. A method according to claim 44, wherein the binding agent is an antibody.

46. A method according to claim 45, wherein the antibody is a monoclonal antibody.

47. A method according to claim 44, wherein the cancer is a prostate cancer.

48. A method for determining the presence or absence of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a prostate-specific protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824, or a complement of any of the foregoing polynucleotide sequences;

(b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; and

(c) comparing the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

49. A method according to claim 48, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a polymerase chain reaction.

50. A method according to claim 48, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a hybridization assay.

51. A method for monitoring the progression of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a prostate-specific protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 1-111, 115-171, 173-175, 177, 179-305, 307-315, 326, 328, 330, 332-335, 340-375, 381, 382 and 384-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-705, 709-774, 777, 789, 817, 823 and 824, or a complement of any of the foregoing polynucleotide sequences;

(b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide;

(c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and

(d) comparing the amount of polynucleotide detected in step (c) to the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

52. A method according to claim 51, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a polymerase chain reaction.

53. A method according to claim 51, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a hybridization assay.

54. A diagnostic kit, comprising:

- (a) one or more antibodies according to claim 11; and
- (b) a detection reagent comprising a reporter group.

55. A kit according to claim 54, wherein the antibodies are immobilized on a solid support.

56. A kit according to claim 54, wherein the detection reagent comprises an anti-immunoglobulin, protein G, protein A or lectin.

57. A kit according to claim 54, wherein the reporter group is selected from the group consisting of radioisotopes, fluorescent groups, luminescent groups, enzymes, biotin and dye particles.

58. An oligonucleotide comprising 10 to 40 contiguous nucleotides that hybridize under moderately stringent conditions to a polynucleotide that encodes a prostate-specific protein, wherein the protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326, 328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772, 779, 817, 823 and 824, or a complement of any of the foregoing polynucleotides.

59. A oligonucleotide according to claim 58, wherein the oligonucleotide comprises 10-40 contiguous nucleotides recited in any one of SEQ ID NO: 2, 3, 8-29, 41-45, 47-52, 54-65, 70, 73-74, 79, 81, 87, 90, 92, 93, 97, 103, 104, 107, 109-111, 115-160, 171, 173-175, 177, 181, 188, 191, 193, 194, 198, 203, 204, 207, 209, 220, 222-225, 227-305, 307-315, 326,

328, 330, 332, 334, 350-361, 363-365, 381, 382, 384, 386, 389, 390, 392, 393, 396, 401, 402, 407, 408, 410, 413, 415-419, 422, 426, 427, 432, 434, 435, 442-444, 446, 450, 452, 453, 459-461, 468-471, 472-476, 524, 526, 530, 531, 533, 535, 536, 552, 569-572, 587, 591, 593-606, 618-689, 691-698, 702-705, 709-772, 779, 817, 823 and 824.

60. A diagnostic kit, comprising:

- (a) an oligonucleotide according to claim 59; and
- (b) a diagnostic reagent for use in a polymerase chain reaction or hybridization assay.

61. A fusion protein according to claim 12, wherein the fusion protein comprises an amino acid sequence selected from the group consisting of: SEQ ID NO: 617, 825 and 835.

63. A fusion protein according to claim 12, wherein the fusion protein comprises an amino acid sequence encoded by a sequence selected from the group consisting of: SEQ ID NO: 616, 822 and 834.

COMPOSITIONS AND METHODS FOR THE THERAPY
AND DIAGNOSIS OF PROSTATE CANCER

ABSTRACT OF THE DISCLOSURE

Compositions and methods for the therapy and diagnosis of cancer, such as prostate cancer, are disclosed. Compositions may comprise one or more prostate-specific proteins, immunogenic portions thereof, or polynucleotides that encode such portions. Alternatively, a therapeutic composition may comprise an antigen presenting cell that expresses a prostate-specific protein, or a T cell that is specific for cells expressing such a protein. Such compositions may be used, for example, for the prevention and treatment of diseases such as prostate cancer. Diagnostic methods based on detecting a prostate-specific protein, or mRNA encoding such a protein, in a sample are also provided.

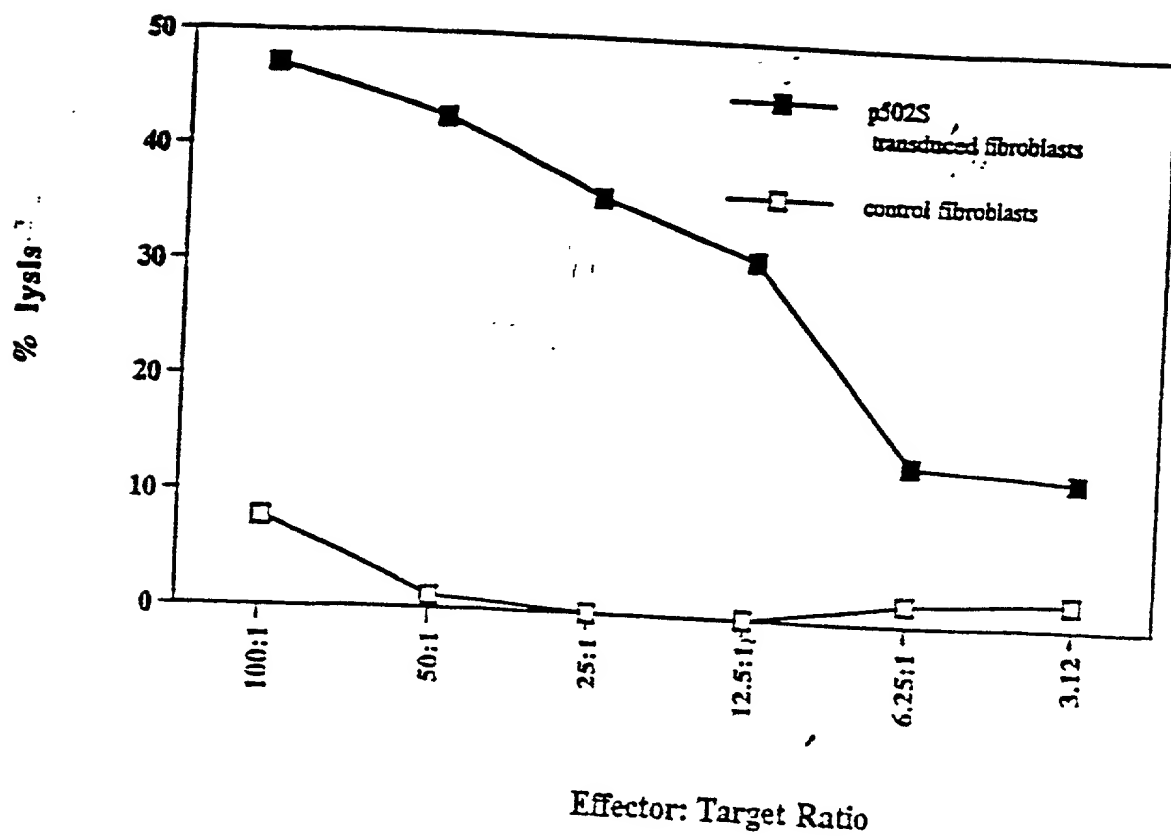


FIG. 1

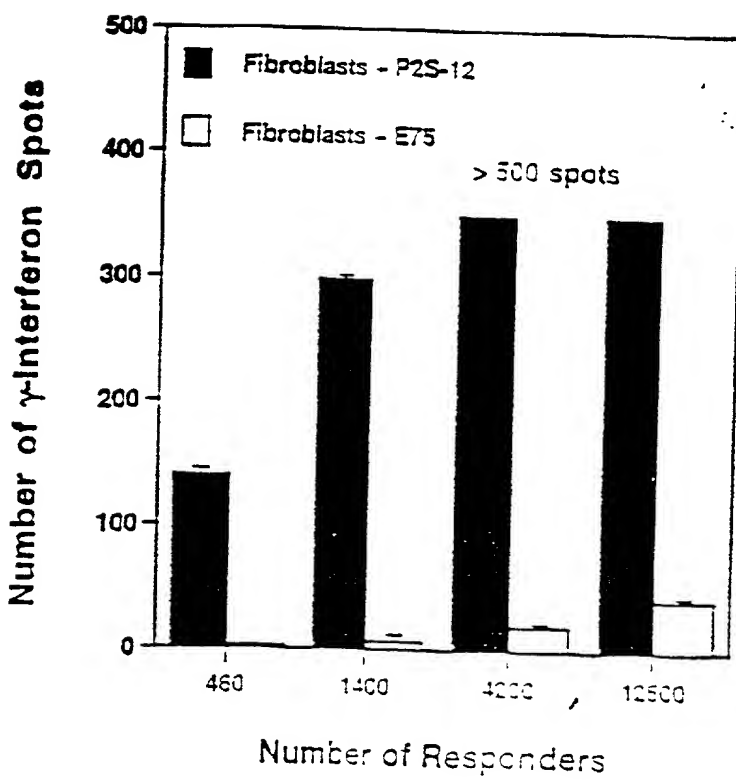


FIG. 2A

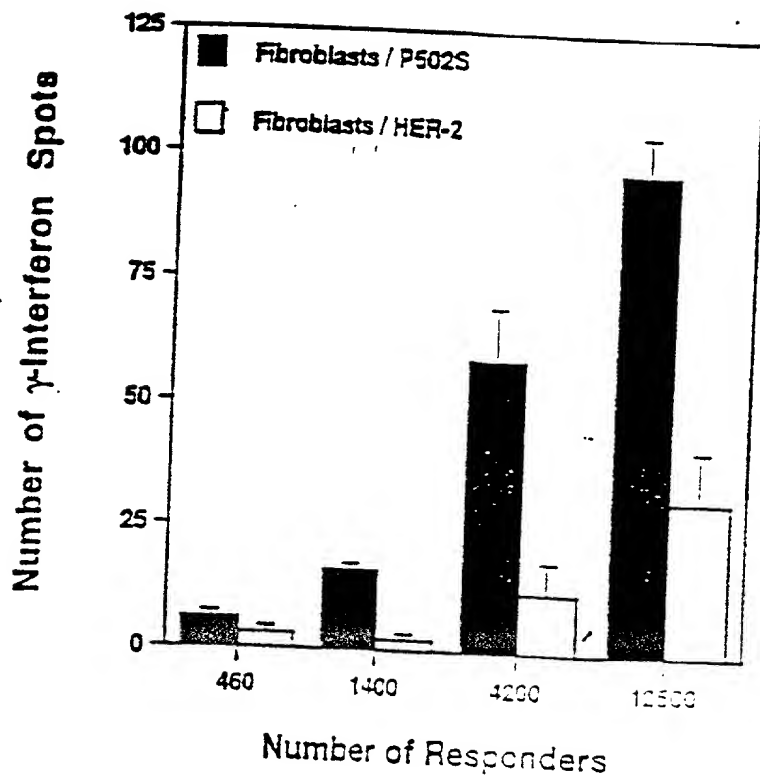


FIG. 2B

002290-28/50960

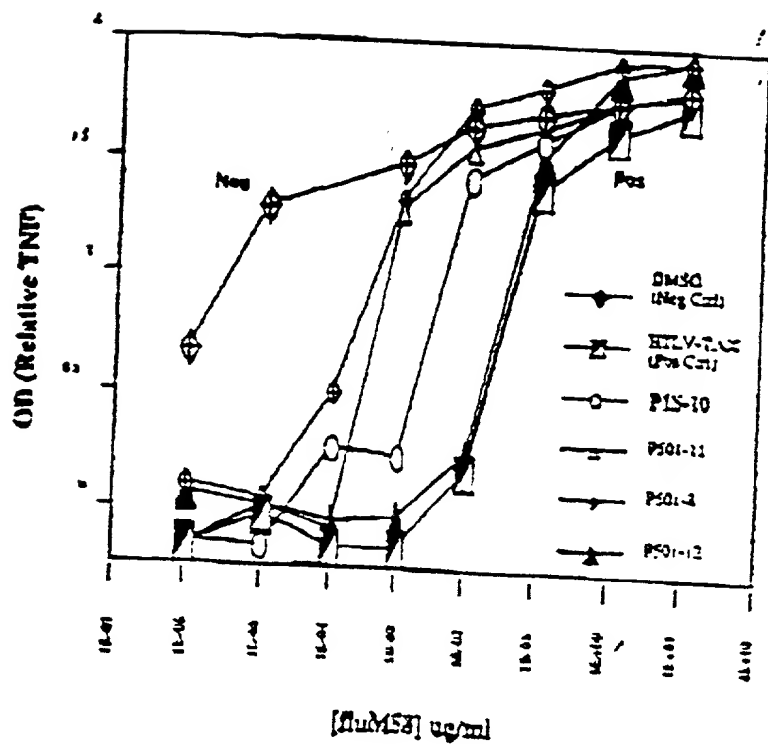


Figure 3

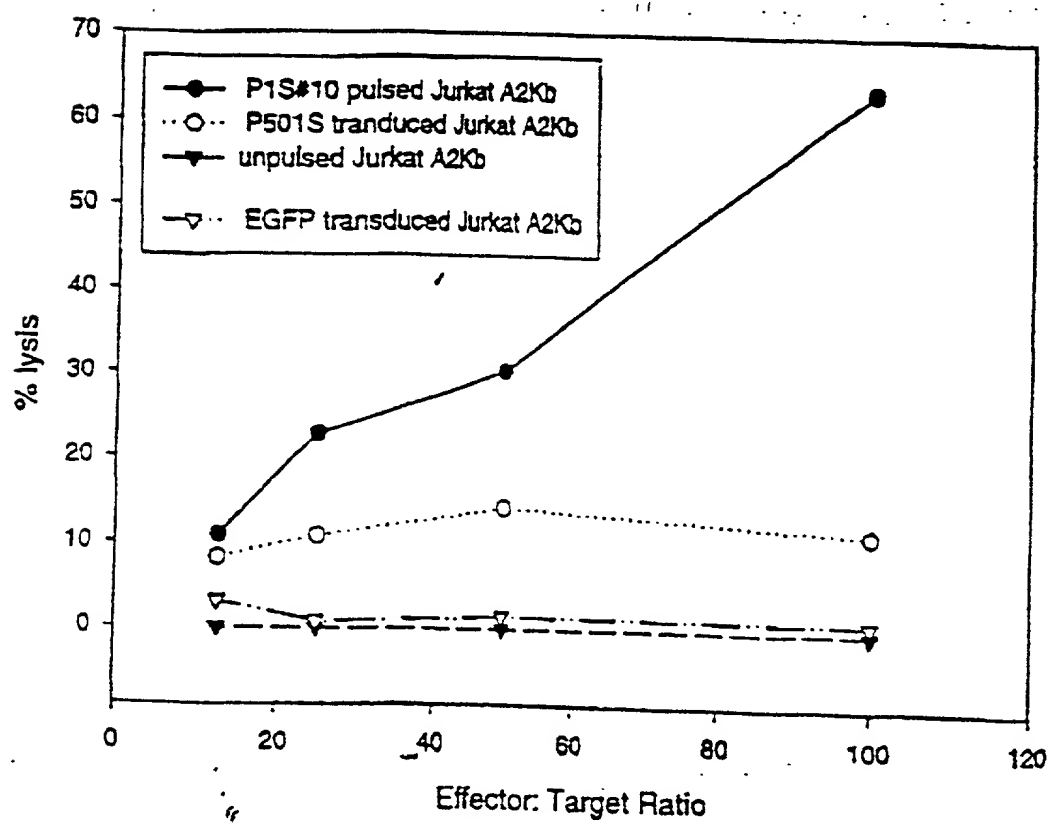


Figure 4

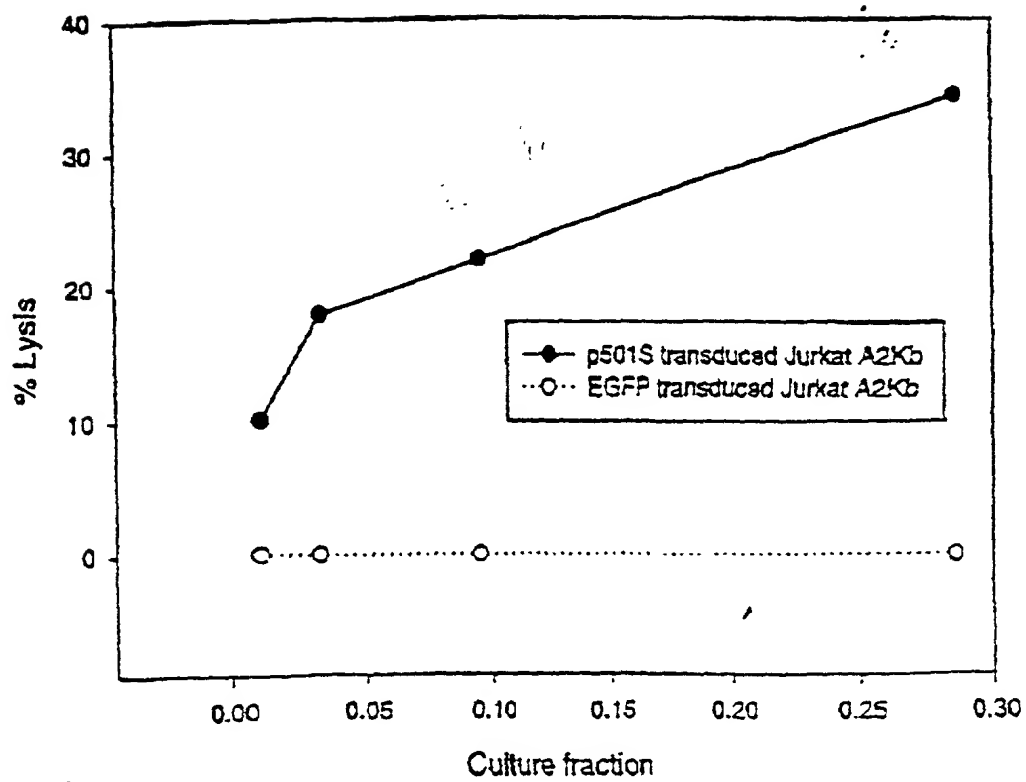


Figure 5

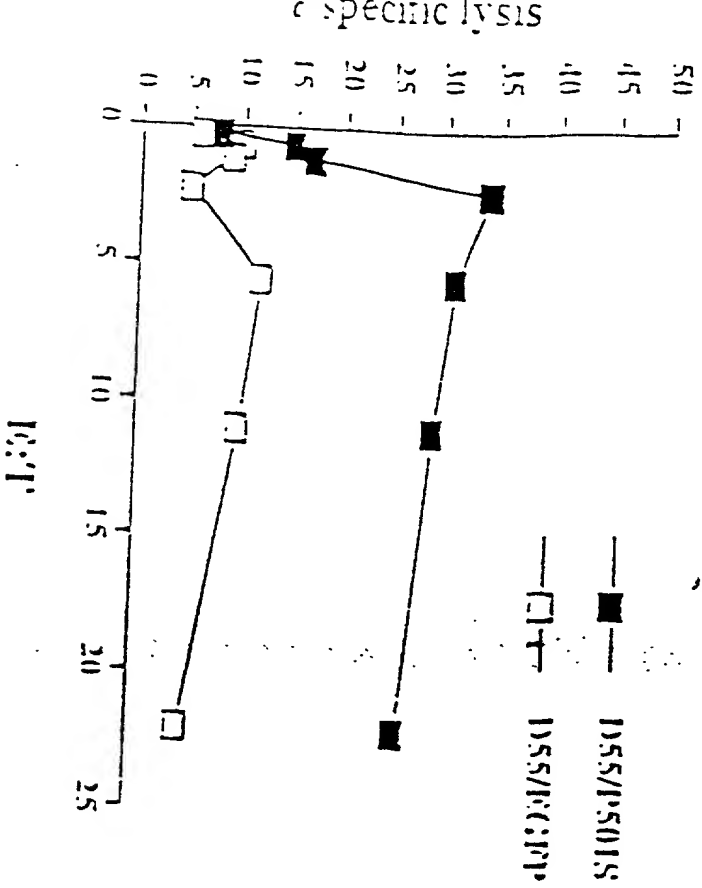


Fig. 6A

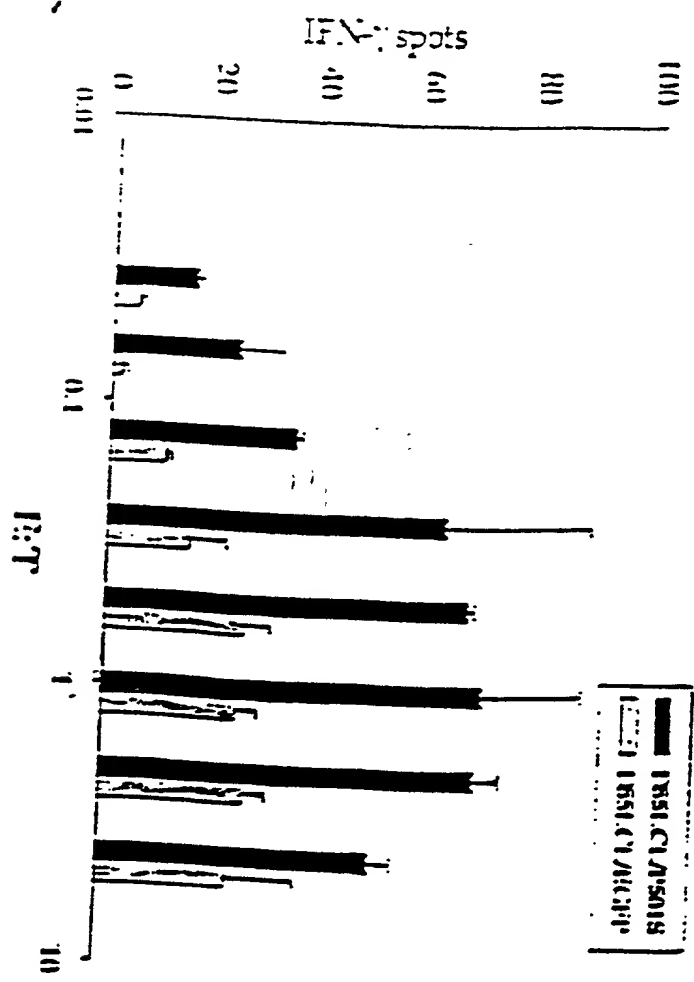
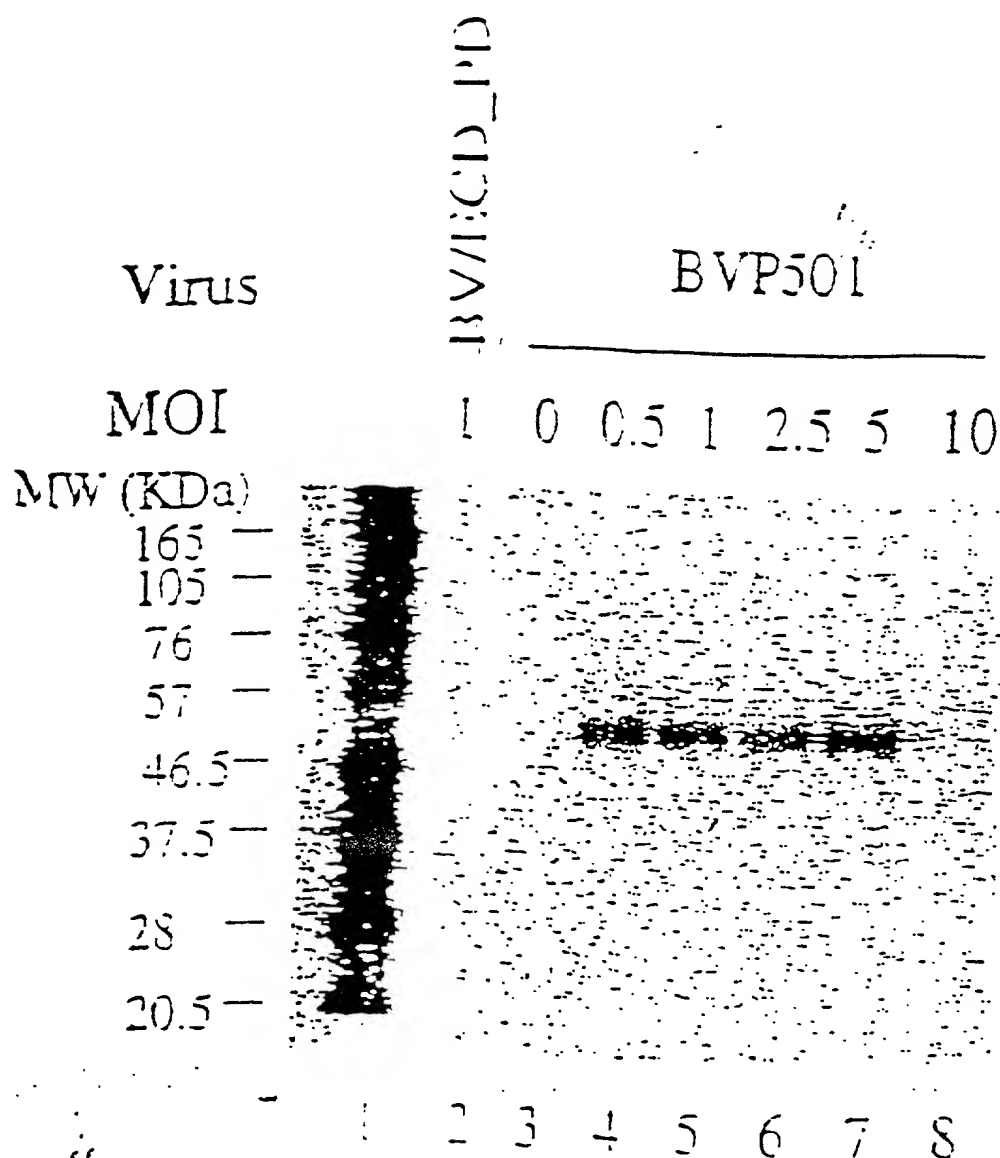


Fig. 6B

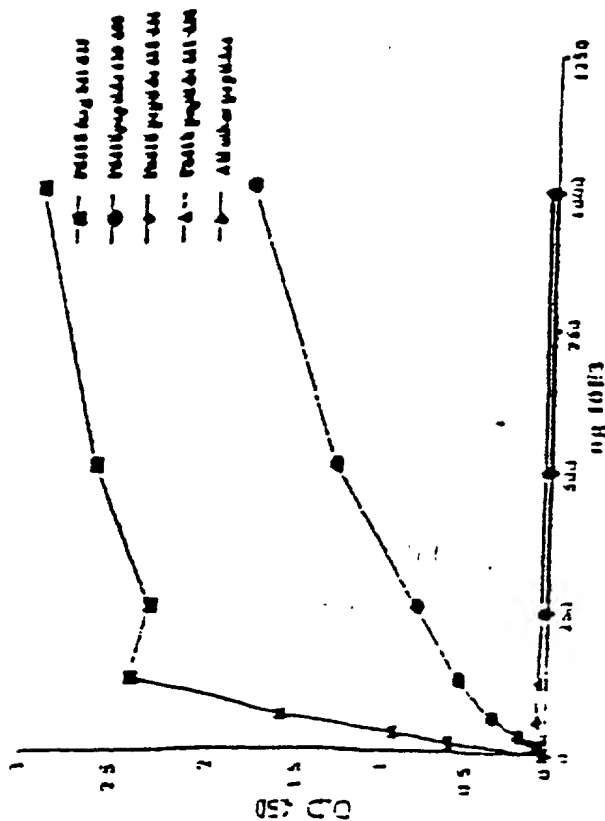
Expression of P501S by the Baculovirus Expression System



0.6 million high titer virus (10⁶ pfu/ml) was added to each well of a 6-well plate were infected with an unrelated control virus BV/ECOD_PD (Lane 2) or with recombinant baculovirus for P501S at different MOI (Lane 3-8). Cell lysates were run on SDS-PAGE under the reducing condition and analyzed by Western blot with a monoclonal antibody against P501S-10E11-G4-D3. Lane 1 is the biotinylated protein molecular weight marker.

Fig 7

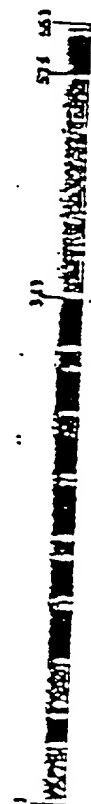
Figure 8. Mapping of the epitope recognized by 10E3-C4-D3



10E3-C4-D3

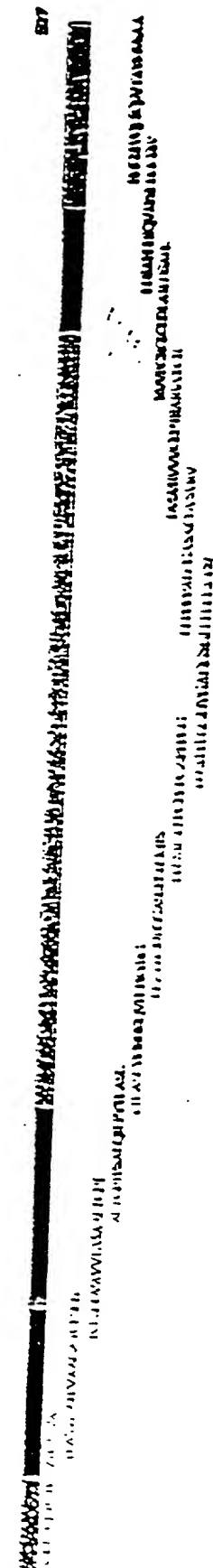
10E3-C4-D3

Full length PSN18



PSN18 fragment used for immunization

10E3-C4-D3



7

Figure 1. Schematic of P501S with predicted transmembrane, cytoplasmic, and extracellular regions

ANVQRLAVVSHLRK AQLALYNLTFFGLVCLAAQT VVPRLLLEVGVEFEEM TMVLGICPVLGLYCVPLLGAS
 DWVRGRYCHRRP FIVALSGLLSLFLPRAGWL AGLLCDDPRPLF LALLILNGLDIFCQNCPTPL
 FALSLDLFRDPDHCRQ AYSYVAFMISLGGCTGYTFPM DWVDSALAPVLCSTQNE
 CLPGLFLFLTCYVAATLY AEFVAGLPPEAEGLSAPVSPHCTPRARAFRMGLALPRL
 DDLCTHAPRTLR LRVAFLLGWMALNTFTTFTYTP VGEELLYQGVYPLAPQTLEARRHYDEGVH
 NQSLQLFLQCAISLYSLVM DRLVQREFCTRAVYLAS VAAFPVAAGIATCLSHSYAVVTA SAA
 LTGFTTSALQLFYTLASLY HREKQVFLPKYRGDTGASSEDSLMTSELPQPKPGAPFPNGHIVQAGCSGL
 LPPPPALCGASACDVSVRVVVEPTFAVVPQERG LCLDLALLDSAPLLSQVAPSLF MGRIVQLSQS
 VTAYMYSAAGILGLVAIYPAT QVVFDKSLAKTSA

Underlined sequence: Predicted transmembrane domain; Bold sequence: Predicted extracellular domain;
 Italic sequence: Predicted intracellular domain. Sequence in bold/underlined: used to generate polyclonal rabbit serum
 Localization of domains predicted using IMMTOPI (G.P. Tusnady and I. Simon (1998) Principles
 Governing Amino Acid Composition of Integral Membrane Proteins: Applications to topology Prediction. J.Mol Biol. 287,
 489-506.

Fig. 9

Genomic Map of (5) Corlca Candidate Genes

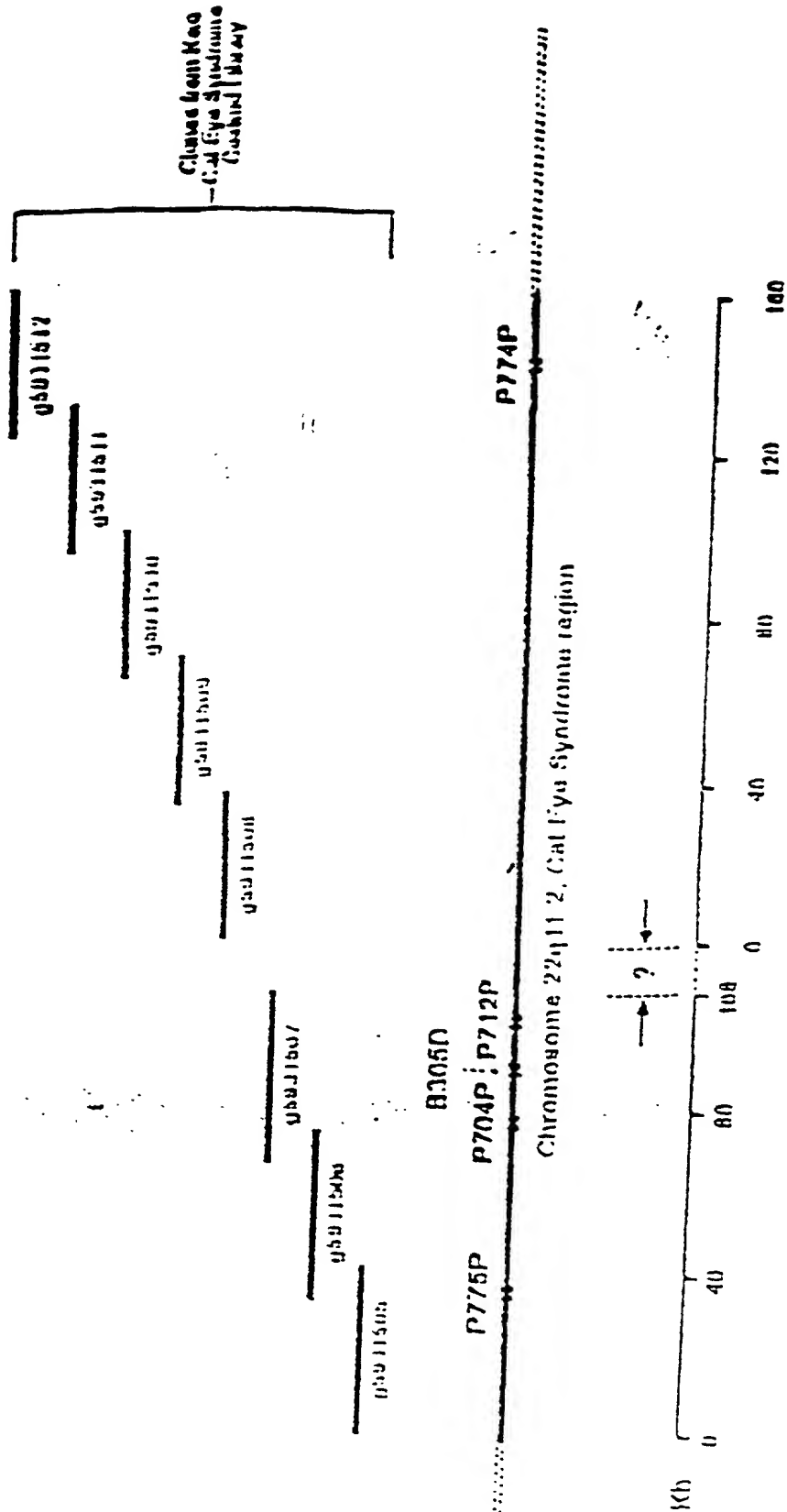
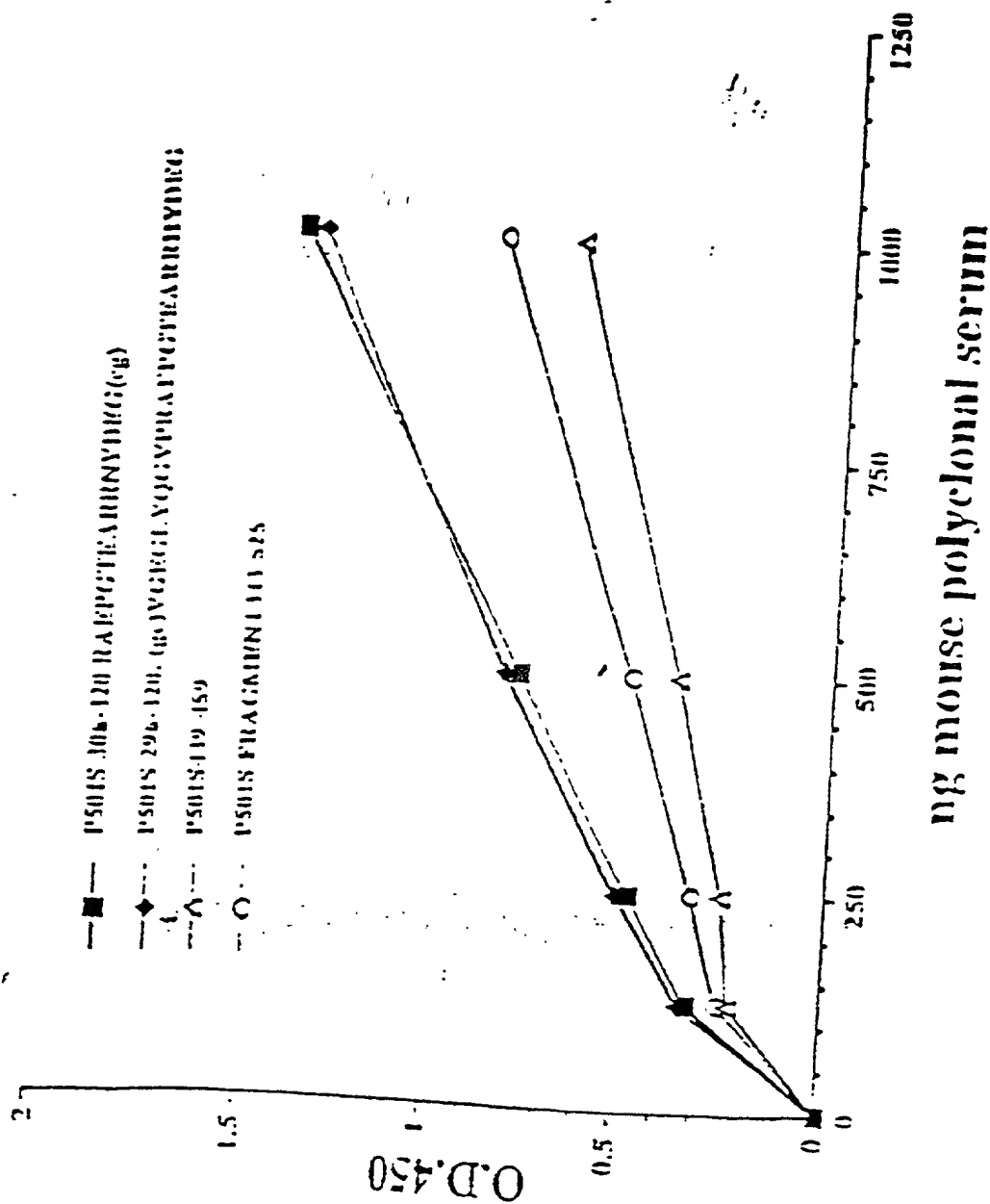


Fig. 10

FIGURE 4. Elisa assay of rabbit polyclonal antibody specificity



10	20	30	40	50	60	70
.....						
GTCAC	TTAGG	AAAGG	TGCTTT	TCGGG	CAGCG	GGGCTCAGCATGAGGAACAGAAAGGAATGACACTCTGG 70
ACAGC	ACCCG	GACCC	TGTACT	CCAGC	CGCTCT	CGGAGCACAGACTTGTCTTACACTGAAAGCGACTTGGT 140
GAATTT	TATTC	AAGCAA	ATTTTA	AGAAAC	AGAAAT	TGTGTCTTCTTTACCAAAGATTCCAAGGCCACGGAG 210
AATGT	GTGCA	AGTGT	GGCTAT	GCCCA	GAGGCC	CAGCAGATGGAAGGCACCCAGATCAACCAAAGTGAGAAAT 280
GGAACT	ACAAG	AAACAC	ACCAAG	GAATTT	CTTACC	BACGCCCTTTGGGGATATTTCAGTTTGAGACACTGGG 350
360	370	380	390	400	410	420
.....						
GAAGAA	AGGGA	AGTATA	TACGTCT	GTCTTC	GCACAC	CGACCGGAAATCCTTTACGAGCTGCTGACCCAG 420
CACTGG	CACCT	GAAAA	CAACCA	ACCTGG	TCATTT	CTGTGACCGGGGGGGCCCAAGAACTTCGCCCTGAAGC 490
CGCGCA	TGCGCA	AGATCT	TCAGCC	GGGCTC	ATCTAC	ATCGCGCAGTCCAAAGGTGCTTGGATTCTCACGGG 560
AGGCAC	CCATT	TGGCT	GACGAA	GTACAT	CGGGG	AGGTGGTGAGAGATAACACTATCAGCAGGAGTTCA 630
GAGGAG	AATATT	TGTGG	CCATT	TGGCAT	AGCAG	CTTGGGGCATGGTCTCCAAACGGGACACCCCTCATCAGGA 700
710	720	730	740	750	760	770
.....						
ATTGGG	ATGCT	GAGGG	CTATTT	TTTAG	CCAGT	ACCTTAAGATGACTTCACAAGGBATCCACTGTATAT 770
CCTGG	ACAAC	CAACAC	ACATTT	TGGT	GGCTGG	TGATGGATATCCCACTGTGCAAGCA 840
AAGCT	CCGGA	ATCAG	CTAGAG	AAGCAT	ATCTCT	GACGGCACTATTCAAGATTCCAACTATGGTGGCAAGA 910
TCCCC	ATTGT	TGTGT	TTTGG	CCAA	AGGAGG	TGAAAAAGAGCTTTGAAAGCCATCAATAGCTCCATCAAAA 980
TAAAA	ATTC	CTTG	TGTGG	TGGT	TGGA	AGGCTCGGGCTGGATCGCTGATGTGATGCTAGCCCTGGTGGAGTG 1050
1060	1070	1080	1090	1100	1110	1120
.....						
GAGGAT	GC	CCG	BACAT	CTTT	CTCG	GTCAAGGAGAGTGGTGGCTTTTACCCCGCACGGTGTCGGG 1120
TGTCT	GAGG	AGGAG	ACTG	AGAG	TTGG	ATCAAAATGGCTCAAGAAATTTCTGAAATGTTCTCACCTATTAA 1190
AGTTA	TTAA	ATGGA	AGAA	CTCG	GGGAT	GAAATTTGAGCAATGGCATCTCTACGGCTCTATACAAAGCC 1260
TTTCA	GCAC	CAAG	TGAG	CAAG	CAAG	ATAACTGGAATGGGC-GTTGAAGTTCTGTGGATGGAAATCAGC 1330
GTG	ACTT	AGCC	AATG	ATGAG	ATTT	TCACCAATGACCGCCATGGAGTCTGCTGACCTTCAAGAAATCAT 1400
1410	1420	1430	1440	1450	1460	1470
.....						
GTTTAC	GGCT	CTCT	CATAAA	AGGAC	AGAC	CCAAAGTTTGTCCGCTCTTTCTGGAGAAATGGCTTGAACCTACGG 1470
AAGTTT	CTC	ACCC	ATGAG	GTCT	CACT	GAACCTCTCTCCAACTACTTCAGCACGCTTGTGTACCGGAATC 1540
TGCAG	ATCG	CCAA	AGAA	TTCT	CTATA	AATGATGCCCTCTCTACGTTTGTCTGGAACCTGGTTGCGAATCTCCG 1610
AAGAG	GGCT	TCGG	AAAG	GAAG	ACAG	AAATGGCTGGGAGAGAGATGGACATAGAACTCCAGGAAGTGTCTCT 1680
ATTACT	CGGC	ACCC	CTCT	GAAG	CTCT	CTCTCATCTGGGCCATCTTTCAGAAAGAAAGGAACCTCTCCAAAG 1750
1760	1770	1780	1790	1800	1810	1820
.....						
TCATTT	GGG	AGCA	BACCA	AGGG	GTG	CACTCTGGCAGCCCTCCGAGCCAGCAAGCTTCTGAAGACTCTGGC 1820
CAAAG	TGA	AGAT	CGAC	ATCA	ATG	CTGTCTGGGAGTCCGAGAGCTGGCTAATGAGTACGAGACCTGGGCT 1890
GTTG	AGCT	GTCT	CACT	BA	GTGT	TACAGCAGCGATGAAGACTTGGCAGAAACAGTGTGTGGTCTATTCTGTG 1960
AAGCT	TTGG	GTGG	AAAG	CAAT	CTCT	GGAGCTGGGCTTGAAGGCAAGGACCACTTTCACCGGCCACCC 2030
TGGGT	TCAG	AAAT	TTCT	TTCT	TAAG	CAATTTGATGAAGATTTCCCGAGACACCAAGAACTGGAAGATT 2100

Fig. 12A (i)

10	20	30	40	50	60	70
MRNRRTLOSTR	LYSSASRSTOL	SYSESQLVNF	IQANFKKREC	VFTKDSKATEN	VCKCGYAQSQHME	70
GTQINQSEKWN	YKXHTKEFPT	DAF30IQFETL	GKXGKYIRLSC	OTDAEILYELL	TQHWHLKTPNLV	140
GGAKNFALKPR	MRKIFSRLLY	IAQSKGAWILT	GGTHYGLTKY	IGEVVRONTISR	SSSEENIVAIGIA	210
VSNROTILRN	COAEGYFLAQY	LMDOFTROPLY	LDNNHTHLLV	DNGCHGHPTEAK	LRNCLEKHSERT	280
IQDSNYGGKIP	IVCFAQGGGKET	LKAINTSIKNK	PCVYVEGSGRI	ADVIASLVEVED	APTSSAVKEKLV	350
360	370	380	390	400	410	420
RFLPRTVSRL	SEETESWIKWLKE	ILECSHLLTV	KMEZAGDEIYS	NAISYALYKAF	STSEQCKDNWNGQ	420
LKLLLEWNCLO	LANDEIFTNORR	WESAOLQEVMT	ALIKORPKFYRL	FLENGLNLRKFL	THOVLTELF3N	490
HFSTLVYRNLI	AKNSYNOCALL	TFVWKLVANFR	RGRKEDRNGR	QENDIELHGVSP	ITRHPLOALFIWA	560
LQNKKELSKV	IWECTRGC	TALALGASKLL	KTAKVKNQINA	AGESEELANEYE	TRAVELTECYSSO	630
AEQLLVYSCEA	WGGSNCLELAVE	ATQCHFTA	3PGVQNFLSK	QWYGEISROTKN	WKILCLFIIPLV	700
710	720	730	740	750	760	770
VSFRKKRYQKH	KKLLWYYVAF	FTSPFVYFSW	NVVFYIAFLL	LFAYVLLMDF	HSVPHPPPELV	LYSLYFVLF 770
CDEVQWYVNGV	NYFTDLWNVM	OTLGLFYFIAG	IVFRLHSBNK	SSLYSGRYFC	LOYIFTLRLIH	FTV 840
SRNLGPKIML	QRMLIDVFF	FLFLFAYWM	VAFGVARCGI	LRONEQRWRW	IFRSVIYEPYLA	MFGQVPSOV 910
DGTTYDFAHCT	FTGNEKFLC	VELDEHNLPR	FEVETIPLV	CIYMLSTNILL	VNLLVAMFGYTV	GTVCEN 980
NOCVWKFCRY	FLVGEYCSRL	NIPFPFIVFAY	FYFMYKXCF	KCCCKEKNM	ESSVCCFKNE	DNETLAWEGYM 1050
1060	1070	1080	1090	1100	1110	1120
KENYLVKINT	KANDTSEEMR	RFRQLODK	LNCLKELLKE	JANKIK		1096

Fig. 12B

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants : Jiangchun Xu et al.
Filed : June 27, 2000
For : COMPOSITIONS AND METHODS FOR THERAPY AND
DIAGNOSIS OF PROSTATE CANCER

Docket No. : 210121.42716C16

Date : June 27, 2000

Box Patent Application
Assistant Commissioner for Patents
Washington, D.C. 20231

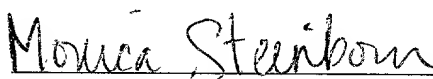
DECLARATION

Sir:

I, Monica Steinborn, in accordance with 37 C.F.R. § 1.821(f) do hereby declare that, to the best of my knowledge, the content of the paper entitled "Sequence Listing" and the computer readable copy contained within the floppy disk are the same.

I declare further that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true and further that these statements are made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Dated this 27th day of June, 2000.



Monica Steinborn
Legal Assistant

701 Fifth Avenue, Suite 6300
Seattle, WA 98104-7092
(206) 622-4900
FAX (206) 682-6031

SEQUENCE LISTING

<110> Xu, Jiangchun
 Dillon, Davin C.
 Mitcham, Jennifer L.
 Harlocker, Susan L.
 Jiang, Yuqui
 Henderson, Robert A.
 Kalos, Michael D.
 Fanger, Gary R.
 Retter, Marc W.
 Stolk, John A.
 Day, Craig H.
 Vedvick, Thomas S.
 Carter, Darrick
 Li, Samuel
 Wang, Aijun
 Skeiky, Yasir A.W.
 Hepler, William

<120> COMPOSITIONS AND METHODS FOR THERAPY AND
 DIAGNOSIS OF PROSTATE CANCER

<130> 210121.42716C16

<140> US

<141> 2000-06-20

<160> 835

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 814

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(814)

<223> n = A,T,C or G

<400> 1

tttttttttt	tttttcacag	tataacagct	ctttatttct	gtgagttcta	ctaggaaatc	60
atcaaatctg	agggttgtct	ggaggacttc	aatacacctc	cccccatagt	gaatcagctt	120
ccaggggggtc	cagtcacctct	ccttacttca	tccccatccc	atgccaaagg	aagaccctcc	180
ctccttgggt	cacagccttc	tctaggcttc	ccagtgcctc	caggacagag	tgggttatgt	240
tttcagctcc	atccttgcgtg	tgagtgtctg	gtgcgttgtg	cctccagctt	ctgctcagtg	300
cttcatggac	agtgtccagc	acatgtcact	ctccactctc	tcagtgtgga	tccactagtt	360
ctagagcggc	cgccaccgcg	gtggagctcc	agcttttgtt	cccttttagtg	agggttaatt	420
gcgcgcttgg	cgtaatcatg	gtcataactg	tttctgtgtg	gaaattgtta	tccgctcaca	480


```

attccacaca acatacgcgc cggaagcata aagtgtaaag cctgggggtgc ctaatgagtg      540
anctaactca cattaattgc gttgcgctca ctgnccgctt tccagtcngg aaaactgtcg      600
tgccagctgc attaatgaat cggccaacgc ncgggggaaaa gcggttttgcg ttttgggggc      660
tcttcgcgctt ctgcgtcaact nantcctgcg ctcggtcntt cggctgcggg gaacgggtatc      720
actcctcaaa ggnnggtatta cggttatccn naaatcnggg gatacccngg aaaaaanttt      780
aacaaaaggg cancaaaggg cngaaacgta aaaa                                814

```

```

<210> 2
<211> 816
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(816)
<223> n = A,T,C or G

```

```

<400> 2
acagaaatgt tggatggtgg agcacctttc tatacgactt acaggacagc agatggggaa      60
ttcatggctg ttggagcaat agaaccccag ttctacgagc tgctgatcaa aggacttgga      120
ctaaagtctg atgaacttcc caatcagatg agcatggatg attggccaga aatgaagaag      180
aagtttgcag atgtatttgc aaagaagacg aaggcagagt ggtgtcaaat ctttgacggc      240
acagatgcct gtgtgactcc ggttctgact tttgaggagg ttgttcatca tgatcacaac      300
aaggaacggg gctcgtttat caccagtgcg gagcaggacg tgagcccccg ccctgcacct      360
ctgctgttaa acaccccagc catcccttct ttcaaaaggg atccactagt tctagaagcg      420
gccgccaccg cgggtggagct ccagcttttg ttccctttag tgagggttaa ttgcgcgctt      480
ggcgtaatca tggatcatagc tgtttcctgt gtgaaattgt tatccgctca caattcccc      540
aacatacgag ccggaacata aagtgttaag cctgggggtgc ctaatgantg agctaactcn      600
cattaattgc gttgcgctca ctgcccgtt tccagtcggg aaaactgtcg tgccactgcn      660
ttantgaatc ngccaccccc cgggaaaagg cgggttgcntt ttgggcctct tccgctttcc      720
tcgctcattg atcctngcnc ccggtcttcg gctgcggnga acggttcact cctcaaaggc      780
ggtntnccgg ttatccccaac acnggggata ccnnga                                816

```

```

<210> 3
<211> 773
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(773)
<223> n = A,T,C or G

```

```

<400> 3
cttttgaaag aagggatggc tgggggtgttt aacagcagag gtgcagggcg ggggctcacg      60
tctgtctcct cactggtgat aaacgagccc cgttccttgt tgtgatcatg atgaacaacc      120
tctcaaaaag tcagaaccgg agtcacacag gcatctgtgc cgtcaaagat ttgacaccac      180
tctgccttcg tcttctttgc aaatacatct gcaaacttct tcttcatttc tggccaatca      240
tccatgctca tctgattggg aagttcacga gactttagtc canntccttt gatcagcagc      300
tcgtagaact ggggttctat tgctccaaca gccatgaatt ccccatctgc tgtcctgtaa      360
gtcgtataga aaggtgctcc accatccaac atgttctgtc ctcgaggggg ggcccgttac      420

```

```

ccaattcgcc ctatantgag tegtattacg cgcgcgcact ggccgctcgtt ttacaacgtc      480
gtgactggga aaaccttggg cgttaccaac ttaatcgctt tgcagcacat ccccttttcg      540
ccagctgggc gtaatancca aaaggcccg cccgatcgcc cttccaacag ttgcgcacct      600
gaatgggnaa atgggacccc cctgttacgg cgcattnaac ccccgcnngg tttngttgtt      660
acccccacnt nnaccgctta cactttgccg gcgccttanc gcccgcctcc tttcnccttt      720
cttccttcc tttcncncn ctttcccccg ggggtttccc cntcaaacc cna                      773

```

```

<210> 4
<211> 828
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(828)
<223> n = A,T,C or G

```

```

<400> 4
cctcctgagt cctactgacc tgtgttttct ggtgtggagt ccagggctgc taggaaaagg      60
aatgggcaga cacagggtga tgccaatgtt tctgaaatgg gtataatttc gtcctctcct      120
tcggaacact ggctgtctct gaagacttct cgctcagttt cagtgaggac acacacaaag      180
acgtgggtga ccatgttggt tgtggggtgc agagatggga ggggtggggc ccacctgga      240
agagtggaca gtgacacaag gtggacactc tctacagatc actgaggata agctggagcc      300
acaatgcatg aggcacacac acagcaagga tgacnctgta aacatagccc acgtgtcct      360
gnnggcactg ggaagcctan atnaggccgt gagcanaaag aaggggagga tccactagtt      420
ctanagcggc cgccaccggg gtgganctcc ancttttggt ccttttagtg agggttaatt      480
gcgcgcttgg cntaatcatg gtcatanctn tttcctgtgt gaaattgtta tccgctcaca      540
attccacaca acatacganc cggaaacata aantgtaaac ctgggggtgc taatgantga      600
ctaactcaca ttaattgcgt tgcgcgcact gcccgccttc caatcnggaa acctgtcttg      660
ccncttgcct tnatgaatcn gccaaacccc ggggaaaagc gtttgcgttt tgggcgctct      720
tccgcttcct cncctantta ntccctnenc tccgtcattc cggctgcngc aaaccggttc      780
accnctcca aaggggggtat tccggtttcc ccnaatccgg gganance                      828

```

```

<210> 5
<211> 834
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(834)
<223> n = A,T,C or G

```

```

<400> 5
tttttttttt tttttactga tagatggaat ttattaagct tttcacatgt gatagcacat      60
agttttaatt gcatccaaag tactaacaaa aactctagca atcaagaatg gcagcatgtt      120
attttataac aatcaacacc tgttggtttt aaaatttggt tttcataaga taatttatac      180
tgaagtaaat ctagccatgc ttttaaaaaa tgcttttaggt cactccaagc ttggcagtta      240
acatttgcca taaacaataa taaaacaatc acaatttaat aaataacaaa tacaacattg      300
taggccataa tcatatacag tataaggaaa aggtggtagt gttgagtaag cagttattag      360
aatagaatac cttggcctct atgcaaatat gtctagacac tttgattcac tcagccctga      420

```

cattcagttt	tcaaagtagg	agacaggttc	tacagtatca	ttttacagtt	tccaacacat	480
tgaaaacaag	tagaaaatga	tgagttgatt	tttattaatg	cattacatcc	tcaagagtta	540
tcaccaaccc	ctcagttata	aaaaattttc	aagttatatt	agtcataata	cttgggtgtgc	600
ttattttaaa	ttagtgtctaa	atggattaag	tgaagacaac	aatgggtcccc	taatgtgatt	660
gatattggtc	atttttacca	gcttctaaat	ctnaactttc	aggcttttga	actggaacat	720
tgnatnacag	tgttccanag	ttncaaccta	ctggaacatt	acagtgtgct	tgattcaaaa	780
tgttattttg	ttaaaaatta	aattttaacc	tggtggaaaa	ataatttgaa	atna	834

<210> 6

<211> 818

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(818)

<223> n = A,T,C or G

<400> 6

tttttttttt	tttttttttt	aagaccctca	tcaatagatg	gagacataca	gaaatagtca	60
aaccacatct	acaaaatgcc	agtatcaggg	ggcggcttcg	aagccaaagt	gatgtttgga	120
tgtaaagtga	aatattagtt	ggcggatgaa	gcagatagtg	aggaaagtgt	agccaataat	180
gacgtgaagt	ccgtggaagc	ctgtggctac	aaaaaatgtt	gagccgtaga	tgccgtcggg	240
aatgggtgaag	ggagactcga	agtactctga	ggcttgtagg	agggtaaaaat	agagaccag	300
taaaattgta	ataagcagtg	cttgaattat	ttggtttcgg	ttgttttcta	ttagactatg	360
gtgagctcag	gtgattgata	ctcctgatgc	gagtaatacg	gatgtgttta	ggagtgggac	420
ttctagggga	tttagcgggg	tgatgcctgt	tggggggccag	tgccctccta	gttgggggggt	480
aggggctagg	ctggagtggg	aaaaggctca	gaaaaatcct	gcgaagaaaa	aaacttctga	540
ggtaataaat	aggattatcc	cgtatcgaag	gccttttttg	acagggtggg	tgtggtggcc	600
ttggtatgtg	ctttctcgtg	ttacatcgcg	ccatcattgg	tatatgggta	gtgtgttggg	660
ttantanggc	ctantatgaa	gaacttttgg	antggaatta	aatcaatngc	ttggccggaa	720
gtcattanga	nggctnaaaa	ggccctgtta	nggggtctgg	ctnggtttta	cccnacccat	780
ggaatncncc	ccccggacna	ntgnatccct	attcttaa			818

<210> 7

<211> 817

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(817)

<223> n = A,T,C or G

<400> 7

tttttttttt	tttttttttt	tggctctaga	gggggtagag	ggggtgctat	agggtaaata	60
cggggccctat	ttcaaagatt	tttaggggaa	ttaattctag	gacgatgggt	atgaaactgt	120
ggtttgctcc	acagatttca	gagcattgac	cgtagtatac	ccccggtcgt	gtagcgtgta	180
aagtggtttg	gttttagacgt	ccgggaattg	catctgtttt	taagcctaata	gtggggacag	240
ctcatgagtg	caagacgtct	tgtgatgtaa	ttattatacn	aatgggggct	tcaatcggga	300
gtactactcg	attgtcaacg	tcaaggagtc	gcaggctcgcc	tggttctagg	aataatgggg	360

gaagtatgta	ggaattgaag	attaatccgc	cgtagtcggt	gttctcctag	gttcaatacc	420
attggtggcc	aattgatttg	atggtaagg	gagggatcgt	tgaactcgtc	tggtatgtaa	480
aggatncctt	ngggatggga	aggcnatnaa	ggactangga	tnaatggcgg	gcangatatt	540
tcaaacngtc	tctanttcct	gaaacgtctg	aaatgttaat	aanaattaan	tttngttatt	600
gaatnttnng	gaaaagggct	tacaggacta	gaaaccaa	angaaaanta	atnntaangg	660
cnttatcntn	aaaggtmata	accnctccta	tnatcccacc	caatngnatt	ccccacncnn	720
acnattggat	nccccanttc	canaaanggc	cnccecccg	tgnannccnc	cttttgttcc	780
cttnantgan	ggttattcnc	ccctngcntt	atcance			817

<210> 8

<211> 799

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1) ... (799)

<223> n = A,T,C or G

<400> 8

catttccggg	tttactttct	aaggaaagcc	gagcgggaagc	tgctaacgtg	ggaatcggtg	60
cataaggaga	actttctgct	ggcacgcgct	agggacaagc	gggagagcga	ctccgagcgt	120
ctgaagcgca	cgtcccagaa	ggtggacttg	gcactgaaac	agctgggaca	catccgcgag	180
tacgaacagc	gcctgaaagt	gctggagcgg	gaggtccagc	agtgtagccg	cgtcctgggg	240
tgggtggccg	angcctganc	cgtctgcct	tgctgcccc	angtgggccc	ccacccccctg	300
acctgcctgg	gtccaaacac	tgagccctgc	tggcggactt	caagganaac	ccccacangg	360
ggattttgct	cctanantaa	ggctcatctg	ggcctcggcc	ccccacctg	gttggccttg	420
tctttgangt	gagccccatg	tccatctggg	ccactgtcng	gaccaccttt	ngggagtgtt	480
ctccttacaa	ccacannatg	ccgggtcct	cccggaaacc	antcccance	tgngaaggat	540
caagnctgn	atccactnnt	netanaaccg	gcncncncg	cngtgggaacc	cnccttntgt	600
tccttttct	tnagggttaa	tnnccgcttg	gccttnccan	ngtcctncnc	nttttccnnt	660
gttnaaattg	ttangcnc	nccnntccn	cnnncnnan	cccgaaccnn	anntnnann	720
ncctgggggt	nccnncgat	tgaccenncc	nccctntant	tgcnttnggg	nncnntgccc	780
ctttccctct	nggganncg					799

<210> 9

<211> 801

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1) ... (801)

<223> n = A,T,C or G

<400> 9

acgccttgat	cctcccaggc	tgggactggt	tctgggagga	gccgggcatg	ctgtggtttg	60
taangatgac	actcccaaag	gtggctcctga	cagtggccca	gatggacatg	gggctcacct	120
caaggacaag	gccaccaggt	gcggggggccg	aagcccacat	gacccctact	ctatgagcaa	180
aatcccctgt	gggggcttct	ccttgaagtc	cgccancagg	gctcagtctt	tggacccang	240
caggctcatg	ggttgtnngc	caactggggg	ccncaacgca	aaanggcnc	gggcctcngn	300

```

caccatcccc angacgcggc tacactnctg gacctccnc tccaccactt tcatgcgctg      360
ttcntaccgc cgnatntgtc ccanctgttt cngtgcenac tccancttct nggacgtgcg      420
ctacatacgc cccgantcnc nctcccgtt tgteccatc cagtnccan caacaaattt      480
cncctantg caccnattcc cacnttttnc agntttccnc nncgngcttc cttntaaaag      540
ggttganccc cggaaaatnc cccaaagggg gggggccngg tacccaactn ccccctnata      600
gctgaantcc ccatnaccnn gnetcnatgg anccntcent tttaannacn ttctnaactt      660
gggaanancc ctgcncntn ccccnnttaa tccnccttg cnangnnct ccccnntcc      720
nccnnntng gcntntnann cnaaaaaggc ccmnnancaa tctcctnncn cctcanttcg      780
ccanccctcg aaatcgccn c

```

<210> 10

<211> 789

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(789)

<223> n = A,T,C or G

<400> 10

```

cagtctatnt ggccagtgtg gcagctttcc ctgtggctgc cggtgccaca tgccctgtccc      60
acagtgtggc cgtggtgaca gttcagccg cctcaccgg gttcaccttc tcagccctgc      120
agatcctgcc ctacacactg gctccctct accaccggga gaagcagggtg ttccctgcca      180
aataccgagg ggacactgga ggtgctagca gtgaggacag cctgatgacc agcttcctgc      240
caggccctaa gcctggagct ccttcccta atggacacgt ggggtgctgga ggcagtggcc      300
tgctcccacc tccaccgcg ctctgogggg cctctgcctg tgatgtctcc gtacgtgtgg      360
tggtgggtga gccaccgan gccagggtgg ttccgggccc gggcatctgc ctggacctcg      420
ccatcctgga tagtgcttcc tgctgtccca ngtggcccca tccctgttta tgggtccat      480
tgtccagctc agccagtctg tcaactgcta tatggtgtct gccgcaggcc tgggtctggt      540
cccatttact ttgtacaca ggtantattt gacaagaacg anttggccaa atactcagcg      600
ttaaaaaatt ccagcaacat tgggggtgga aggcctgcct cactgggtcc aactccccgc      660
tcctgttaac ccatggggc tgccggcttg gccgccatt tctgttgctg ccaaantnat      720
gtggctctct gctgccacct gttgctggct gaagtgcnta cngcncanct nggggggtng      780
gngttccc

```

<210> 11

<211> 772

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(772)

<223> n = A,T,C or G

<400> 11

```

cccaccctac ccaaataatta gacaccaaca cagaaaagct agcaatggat tcccttctac      60
ttgtttaat aaataagtta aatattttaa tgccgtgtgc tctgtgatgg caacagaagg      120
accaacaggc cacatcctga taaaaggtaa gaggggggtg gatcagcaaa aagacagtgc      180
tgtgggctga ggggacctgg ttcttgtgtg ttgcccctca ggactcttcc cctacaaata      240

```

```

actttcatat gttcaaatcc catggaggag tgtttcatcc tagaaactcc catgcaagag      300
ctacattaaa cgaagctgca ggtaagggg cttanagatg ggaaaccagg tgactgagtt      360
tattcagctc caaaaaaccc ttctctaggt gtgtctcaac taggaggcta gctgttaacc      420
ctgagcctgg gtaatccacc tgcagagtc cgcattcca gtgcatggaa ccttctggc      480
ctccctgtat aagtccagac tgaaaccccc ttggaaggnc tccagtcagg cagccctana      540
aactggggaa aaaagaaaag gacgccccan ccccgagctg tgcanctacg cacctcaaca      600
gcacagggtg gcagcaaaaa aaccacttta ctttggcaca aacaaaaact nggggggggca      660
accccggcac cccnangggg gttaacagga ancngggnaa cntggaaccc aattnaggca      720
ggccnccac cccnaatntt gctgggaaat ttttctccc ctaaattntt tc              772

```

<210> 12

<211> 751

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(751)

<223> n = A,T,C or G

<400> 12

```

gccccaatc cagctgccac accacccacg gtgactgcat tagttcggat gtcatacaaa      60
agctgattga agcaaccctc tactttttgg tcgtgagcct tttgcttggg gcaggtttca      120
ttggctgtgt tggtgacgtt gtcattgcaa cagaatgggg gaaaggcact gttctctttg      180
aagtanggtg agtcctcaaa atccgtatag ttgggtgaagc cacagcactt gagccctttc      240
atgggtggtg tccacacttg agtgaagtct tccgtgggaa cataatcttt cttgatggca      300
ggcactacca gcaacgtcag ggaagtgtct agccattgtg gtgtacacca aggcgaccac      360
agcagctgcn acctcagcaa tgaagatgan gaggangatg aagaagaacg tcncgagggc      420
acacttgctc tcagtcttan caccatanca gcccntgaaa accaananca aagaccacna      480
cnccggctgc gatgaagaaa tnaccccneg ttgacaaact tgcatggcac tggganccac      540
agtggccna aaaatcttca aaaaggatgc cccatcnatt gaccccccaa atgccactg      600
ccaacagggg ctgccccacn cncnnaacga tganccnatt gnacaagatc tncntggtct      660
tnatnaacnt gaaccctgcn tngtggctcc tgttcaggnc cnnggcctga cttctnaann      720
aangaactcn gaagncccca cngganannc g                                751

```

<210> 13

<211> 729

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(729)

<223> n = A,T,C or G

<400> 13

```

gagccaggcg tccctctgcc tgcccactca gtggcaacac ccgggagctg ttttgcctt      60
tgtggancct cagcagtncc ctctttcaga actcantgcc aaganccctg aacaggagcc      120
accatgcagt gcttcagctt cattaagacc atgatgatcc tcttcaattt gctcatcttt      180
ctgtgtggtg cagccctgtt ggcagtgggc atctgggtgt caatcgatgg ggcacacctt      240
ctgaagatct tggggccact gtcgtccagt gccatgcagt ttgtcaacgt gggctacttc      300

```

```

ctcatcgag cccggtgtgt ggtcttagct ctaggtttcc tgggctgcta tgggtgctaag 360
actgagagca agtgtgccct cgtgacgttc ttcttcatcc tctctctcat cttcattgct 420
gaggttgcaa tgctgtggtc gccttggtgt acaccacaat ggctgagcac ttcttgacgt 480
tgctggtaat gcctgccatc aanaaaagat tatgggttcc caggaanact tcaactcaagt 540
gttgaacac caccatgaaa gggctcaagt gctgtggctt cnnccaacta tacggatttt 600
gaagantcac ctacttcaaa gaaaanagtg cctttccccc atttctgttg caattgacaa 660
acgtcccaaa cacagccaat tgaaaacctg caccacaacc aaanggggtcc ccaaccanaa 720
attnaaggg

```

```

<210> 14
<211> 816
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(816)
<223> n = A,T,C or G

```

```

<400> 14
tgctcttcct caaagttggt cttgttgcca taacaaccac cataggtaaa gcggggcgag 60
tgctcgctga aggggttgta gtaccagcgc gggatgctct ccttgagag tctgtgtct 120
ggcaggtcca cgcagtgcc tttgtcactg gggaaatgga tgcgctggag ctctgcaaag 180
ccactcgtgt atttttcaca ggcagcctcg tccgacgcgt cggggcagtt ggggggtgtct 240
tcacactcca ggaaactgtc natgcagcag ccattgctgc agcggaactg ggtgggctga 300
cangtgccag agcacactgg atgggcgctt tccatgnnan gggccctgng ggaaagtccc 360
tgancccan anctgcctct caaangcccc acctgacaca ccccgacagg ctagaatgga 420
atcttcttcc cgaaaggtag ttnttcttgt tgcccaancc anccccntaa acaaactctt 480
gcanatctgc tccgnggggg tentantacc ancgtgggaa aagaaccca ggcngcgaa 540
caancttggt tggatnccga gcnataatct nctnttctgc ttggtggaca gcaccantna 600
ctgtnnanct ttagncntg gtectentgg gttgnncttg aacctaatn ccnntcaact 660
gggacaaggt aantngcent cctttnaatt cccnancntn cccctggtt tgggggtttt 720
cncnctcta cccagaaan nccgtgttcc cccccaacta ggggcnaaa ccnntnttc 780
cacaaccctn cccacccac gggttcngnt ggttng 816

```

```

<210> 15
<211> 783
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(783)
<223> n = A,T,C or G

```

```

<400> 15
ccaaggcctg ggcaggcata nacttgaagg tacaaccca ggaaccctg gtgctgaagg 60
atgtggaaaa cacagattgg cgcctactgc ggggtgacac ggatgtcagg gtagagagga 120
aagacccaaa ccaggtggaa ctgtggggac tcaaggaang cacctacctg ttccagctga 180
cagtgactag ctcagaccac ccagaggaca cggccaactg cacagtcact gtgctgtcca 240
ccaagcagac agaagactac tgcctcgcat ccaacaangt ggtcgcctgc cggggctctt 300

```

```

tcccacgctg gtactatgac cccacggagc agatctgcaa gagtttcggt tatggaggct 360
gcttgggcaa caagaacaac taccttcggg aagaagagtg cattctancc tgcnggggtg 420
tgcaagggtg gcctttgana ngcanctctg gggctcangc gactttcccc caggggccct 480
ccatggaaaag gcgccatcca ntgttctctg gcacctgtca gcccacccag ttccgctgca 540
ncaatggctg ctgcatcnac antttcctng aattgtgaca acacccccca ntgcccccaa 600
ccctcccaac aaagcttccc tgttnaaaaa tacnccantt ggcttttnac aaacncccg 660
cncctccttt ttcccnntn aacaaagggc nctngccttt gaactgcccn aaccnnggaa 720
tctnccnngg aaaaantncc cccctgggt cctnnaancc cctccncaaa anctncccc 780
ccc 783

```

```

<210> 16
<211> 801
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(801)
<223> n = A,T,C or G

```

```

<400> 16
gccccaatc cagctgccac accacccacg gtgactgcat tagttcggat gtcatacaaaa 60
agctgattga agcaaccctc tacttttttg tctgtagcct tttgcttggg gcaggtttca 120
ttggctgtgt tgggtgacgtt gtcattgcaa cagaatgggg gaaaggcact gttctctttg 180
aagtaggggtg agtcctcaaa atccgtatag ttgggtgaagc cacagcactt gagccctttc 240
atggtggtgt tccacacttg agtgaagtct tcctgggaac cataatcttt cttgatggca 300
ggcactacca gcaacgtcag gaagtgtca gccattgtgg tgtacacca ggcgaccaca 360
gcagctgcaa cctcagcaat gaagatgagg aggaggatga agaagaacgt cncgagggca 420
cacttgctct ccgtcttagc accatagcag cccangaaac caagagcaaa gaccacaacg 480
ccngctgcga atgaaagaaa ntaccacgt tgacaaactg catggccact ggacgacagt 540
tggcccgaa atcttcagaa aagggatgcc ccatcgattg aacaccana tgccactgc 600
cnacagggct gcnccnncn gaaagaatga gccattgaag aaggatcntc ntggctcttaa 660
tgaactgaaa cntgcatgg tggccctgt tcagggctct tggcagtga ttctganaaa 720
aaggaacngc ntnagcccc ccaaangana aaacaccccc ggggtgttgc ctgaattggc 780
ggccaaggan cctgccccn g 801

```

```

<210> 17
<211> 740
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(740)
<223> n = A,T,C or G

```

```

<400> 17
gtgagagcca ggcgtccctc tgccctgccc ctgagtggca acacccggga gctgttttgt 60
cctttgtgga gcctcagcag ttccctcttt cagaactcac tgccaagagc cctgaacagg 120
agccaccatg cagtgttca gcttcattaa gaccatgatg atcctcttca atttgctcat 180
ctttctgtgt ggtgcagccc tgttggcagt gggcatctgg gtgtcaatcg atggggcatc 240

```



```

ctttctgaag atcttcgggc cactgtcgtc cagtgccatg cagtttgtca acgtgggcta      300
cttcctcatc gcagccggcg ttgtggtctt tgctcttggt ttcttgggct gctatgggtgc      360
taagacggag agcaagtgtg cctcgtgac gttcttcttc atcctcctcc tcctcttcat      420
tgctgaagtt gcagctgctg tggtcgcctt ggtgtacacc acaatggctg aaccattcct      480
gacgttgctg gtantgcctg ccatcaanaa agattatggg tcccaggaa aaattcactc      540
aantntggaa caccnccatg aaaagggctc caatttctgn tggcttcccc aactataccg      600
gaattttgaa agantcncct tacttccaaa aaaaaanant tgccttttnc cccnttctgt      660
tgcaatgaaa acntcccaan acngccaatn aaaaactgcc cnnncaaaaa ggntcncaaa      720
caaaaaaant nnaaggggtn

```

```

<210> 18
<211> 802
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(802)
<223> n = A,T,C or G

```

```

<400> 18
ccgctgggtg cgctgggtcca gngnagccac gaagcacgtc agcatacaca gcctcaatca      60
caaggtcttc cagctgccgc acattacgca gggcaagagc ctccagcaac actgcatatg      120
ggatacactt tacttttagca gccaggggtga caactgagag gtgtcgaagc ttattcttct      180
gagcctctgt tagtggagga agattccggg cttcagctaa gtagtcagcg tatgtcccat      240
aagcaaacac tgtgagcagc cggaaggtag aggcaaagtc actctcagcc agctctctaa      300
cattggggcat gtccagcagt tctccaaaca cgtagacacc agnggcctcc agcacctgat      360
ggatgagtgt ggccagcgct gcccccttgg ccgacttggc taggagcaga aattgctcct      420
ggttctgccc tgtcaccttc acttccgcac tcactactgc actgagtgtg ggggacttgg      480
gctcaggatg tccagagacg tggttccgcc cctcncctta atgacaccgn ccanncaacc      540
gtcggctccc gccgantgng ttcgctgtn ctaggtcagg gtctgctggc cnetacttgc      600
aancttcgtc nggccccatg aattcaccnc accggaactn gtangatcca ctntttctat      660
aaccggnctc caccgcnntt ggaactccac tcttnttnc tttacttgag ggtaaggctc      720
acccttnnct ttaccttggt ccaaaccntn cntgtgtcg anatngtnaa tcnggncna      780
tnccancnc atangaagcc ng

```

```

<210> 19
<211> 731
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(731)
<223> n = A,T,C or G

```

```

<400> 19
cnaagcttcc aggtnacggg ccgcnaanc tgaccnagg tancanaang cagnncgctg      60
gagccaccg tcacngngng gngtctttat nggagggggc ggagccacat cnetggacnt      120
cntgacccca actccccnc nncantgca gtgatgagtg cagaactgaa ggtnacgtgg      180
caggaaccaa gancaaannc tgctccnntc caagtcggcn nagggggcgg ggctggccac      240

```

```

gencatccnt cnagtgtctgn aaagccccnn cctgtctact tgtttggaga acngcnnnga      300
catgcccagn gttanataac nggcngagag tnannttgcc tctcccttcc ggctgcgcan      360
cngntntgct tagnggacat aacctgacta cttaactgaa cccnngaate tncncncct      420
ccactaagct cagaacaaaa aacttcgaca ccactcantt gtcacctgnc tgcctcaagta      480
aagtgtaccc catncccaat gtntgtctnga ngctctgncc tgcnttangt tcgggtcctgg      540
gaagacctat caattnaagc tatgtttctg actgcctctt gctccctgna acaancnacc      600
cnncnntcca aggggggggnc ggcccccaat ccccccaacc ntnaattnan tttancccn      660
ccccnnggcc cggcctttta cnancntcnn nnacngggna aaaccnnngc tttncccaac      720
nnaatccncc t                                                    731

```

```

<210> 20
<211> 754
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(754)
<223> n = A,T,C or G

```

```

<400> 20
tttttttttt tttttttttt taaaaacccc ctccattnaa tgnaaacttc cgaaattgtc      60
caacccctc ntccaaatnn cnttttcgg gnggggggttc caaacccaan ttanntttgg      120
annttaaatt aaatnttntt tggnggnnna anccnaatgt nangaaagtt naaccanta      180
tnancctnaa tncctggaaa cngtngntt ccaaaaatnt ttaaccctta antccctccg      240
aaatngttna nggaaaaccc aanttctcnt aagggtgttt gaaggntnaa tnaaaanccc      300
nnccaattgt ttttngccac gcctgaatta attggnnttc gntgttttcc nttaaaanaa      360
ggnnancccc gggtantnaa tcccccnnc cccaattata ccganttttt ttngaattgg      420
ganccncgg gaattaacgg ggnnnnntccc tnttgggggg cnggnncccc cccntcggg      480
ggttngggnc aggnennaat tgtttaaggg tccgaaaaat ccctccnaga aaaaaanctc      540
ccaggntgag nntnggggtt ncccccccc cangggccct ctcgnanagt tgggggtttg      600
ggggcctggg attttnttcc ccctnttnc tcccccccc ccnggganag aggttngngt      660
ttgtntcnnn ggcccccncc aaganccttn ccganttnan ttaaatecnt gcctnggcga      720
agtcctntgn agggntaaan ggccccctnn cggg                                                    754

```

```

<210> 21
<211> 755
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(755)
<223> n = A,T,C or G

```

```

<400> 21
atcancccat gacccenaac nngggaccnc tcancgggnc nnncnaccnc cggccnatca      60
nngtnagnnc actncnnttn natcacnccc cncnactac gccncnanc cnacgcncta      120
nncanatncc actganngcg cgangtngan ngagaaanct nataccanag ncaccanacn      180
ccagctgtcc nanaangcct nnnatacngg nnnateccaat ntgnancctc cnaagtattn      240
nncnncanat gattttcctn anccgattac cntncccccc tanccctcc cccccaacna      300

```



```

ctnccnacc  tacntcttcn  nagctgtcnn  acccctngtn  cgnaccccc  naggtcgggg  300
tcgggttttn  nntgaccgng  cnnccccctc  ccccntccat  nacgancnc  ccgcaccacc  360
nanngcncgc  nccccgnnct  cttegcncnc  ctgtcctntn  cccctgtngc  ctggcncngn  420
accgcattga  ccctcgccnn  ctncnngaaa  ncgnanacgt  ccgggttggn  annancgctg  480
tgggnnnngc  tctgcncgc  gtctcttcn  ncncttcca  ccattctct  tacnggggtc  540
ccncccntc  tcnnncacnc  cctgggacgc  tntcctntgc  ccccttnac  tccccctt  600
cgcgtgnc  cgnccccacc  ntcatttnca  nacgntcttc  acaanncct  ggntnctcc  660
cnancngnc  gtcancnag  ggaagggngg  ggnncnntg  nttgacgttg  ngngangtc  720
cgaanantcc  tcncntcan  cncctaccct  cgggcgnnct  ctngttnc  aacttancaa  780
ntctcccccg  ngngcncntc  tcagcctcnc  cccccnct  ctctgcantg  tntctgctc  840
tnaccnntac  gantnttcgn  cncctcttt  cc  872

```

<210> 24

<211> 815

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(815)

<223> n = A,T,C or G

<400> 24

```

gcatgcaagc  ttgagtattc  tatagngtca  cctaaatanc  ttggcntaat  catggtcnta  60
nctgncttcc  tgtgtcaaag  gtatacnaa  tanatatgaa  tctnatntga  caaganngta  120
tcntncatta  gtaacaantg  tnntgtccat  cctgtengan  canattccca  tnnattncgn  180
cgcattcn  gncantatn  taatngggaa  ntcnnntnnn  ncaccnncat  ctatcntncc  240
gncacctgac  tggagagat  ggatnantt  tnntntgacc  nacatgttca  tcttggattn  300
aanancccc  cgcngnccac  cggttngng  cnagccnntc  ccaagacctc  ctgtggaggt  360
aacctgcgtc  aganncatca  aacntgggaa  accgcnncc  angtnnaagt  ngnnncanan  420
gatcccgacc  aggnntnacc  atcccttcnc  agcgcacct  ttngtgcctt  anagnnagc  480
gtgtccnanc  cncctcaacat  ganaagcgcc  agnccanccg  caattnggca  caatgtcgnc  540
gaacccccct  gggggganna  tncaaanccc  caggattgtc  cncncangaa  atccncanc  600
ccnccctac  ccncttttg  gacngtgacc  aantcccgga  gtncagtc  ggcngnctc  660
ccccaccgt  nncntgggg  ggggtgaant  cngnntcanc  cngncgaggn  ntcgnaagga  720
accgncctn  ggcgaanng  ancnntcnga  agnccnct  cgtataacce  cccctcncca  780
nccnancng  agntcccccc  cngggtnccg  aangg  815

```

<210> 25

<211> 775

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(775)

<223> n = A,T,C or G

<400> 25

```

ccgagatgtc  tgcctcgtg  gccttagctg  tgctcgct  actctctctt  tctggcctgg  60
aggctatcca  gcgtactcca  aagattcagg  tttactcacg  tcatccagca  gagaatggaa  120

```

```

agtcaaattt cctgaattgc tatgtgtctg ggtttcatcc atccgacatt gaanttact 180
tactgaagaa tgganagaga attgaaaaag tggagcattc agacttgtct ttcagcaagg 240
actggctctt ctatctcntg tactacactg aattcacccc cactgaaaaa gatgagtatg 300
cctgccgtgt gaaccatgtg actttgtcac agcccaagat agttaagtgg gatcgagaca 360
tgtaagcagn cnnecatggaa gtttgaagat gccgcatttg gattggatga attccaaatt 420
ctgcttgctt gcnttttaat antgatatgc ntatacacc taccctttat gnccccaaatt 480
tgtaggggtt acatnantgt tcnctnngga catgatcttc ctttataant ccncnttcg 540
aattgccgt cncncngttn ngaatgtttc cnaaaccacg gttggctccc ccaggtcncc 600
tcttacggaa gggcctgggc cnccttncaa ggttggggga accnaaaatt tcncttntgc 660
ccnccncca cnccttgng nncncanttt ggaacccttc cnattcccct tggcctcnna 720
nccttnncta anaaaacttn aaancgtngc naaanntttn acttcccccc ttacc 775

```

<210> 26

<211> 820

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(820)

<223> n = A,T,C or G

<400> 26

```

anattantac agtgtaatct tttcccagag gtgtgtanag ggaacggggc ctagaggcat 60
cccanagata ncttatanca acagtgtctt gaccaagagc tgctgggcac atttcctgca 120
gaaaagggtg cgggtcccat cactcctcct ctcccatagc catcccagag gggtgagtag 180
ccatcangcc ttcgggtggga gggagtcang gaaacaacan accacagagc anacagacca 240
ntgatgacca tgggcgggag cgagcctctt ccctgnaccg gggtggcana nganagccta 300
nctgaggggt cacactataa acgttaacga ccnagatnan cacctgttc aagtgcaccc 360
ttcctacctg acnaccagng accnnnaact gncgcctggg gacagcnctg ggancagcta 420
acnnagcact cacctgcccc cccatggccg tncgntccc tggctcctgnc aagggaagct 480
cctgttgga attncgggga naccaaggga nccccctcct ccantgtga aggaaaaann 540
gatggaattt tnccttccg gccnntcccc tcttcttta cagccccct nntactctc 600
tccctctntt ntectgnenc acttttnacc ccnnnatttc ccttnattga tcggannctn 660
ganattccac tnnegcctnc cntcnatcng naanacnaaa nactntctna ccnnggggat 720
gggnncctcg ntcctcctct ctttttctct accncnntt ctttgctct ccttngatca
780tccaaccntc gntggccntn ccccccnnn tcttttccc 820

```

<210> 27

<211> 818

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(818)

<223> n = A,T,C or G

<400> 27

```

tctgggtgat ggctcttcc tctcagga cctctgactg ctctgggcca aagaatctct 60
tgtttctct cagagcccca ggcagcgggt attcagccct gcccaacctg attctgatga 120

```

```

ctgcggatgc tgtgacggac ccaaggggca aataggggtcc caggggtccag ggagggggcgc      180
ctgctgagca cttccgcccc tcacctgccc cagccctctgc catgagctct gggctgggtc      240
tccgcctcca gggttctgct cttccangca ngccancaag tggcgctggg ccacactggc      300
ttcttcctgc ccntccctg gctctgante tctgtcttcc tgcctgtgc angcnccttg      360
gatctcagtt tccctcnctc anngaactct gtttctgann tcttcantta actntgantt      420
tatnaccnan tggnetgtnc tgtcnnaact taatgggccc gaccggctaa tccctccctc      480
nctcccttcc anttcnnnna accngettnc cntctctccc centancccg ccngggaanc      540
ctcctttgcc ctnaccangg gccnnnaccg cccntnnctn ggggggcnnng gtnnctncnc      600
ctgntnnccc cncctcnctt tncctctgccc cnnnncgcgn nngcannttc ncngtcccn      660
tnnctcttcn ngntnecnaa ngntcnctn tnnnnngncn ngntnntncn tccctctcnc      720
cnnntgnang tnnntnnnnc ncngnncccc nnnnncnnnn nggnntnnnn tctnncncgc      780
cccncccccc ngnattaagg cctccnntct cgggcnc      818

```

```

<210> 28
<211> 731
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(731)
<223> n = A,T,C or G

```

```

<400> 28
aggaagggcg gagggatatt gtanggggatt gagggatagg agnataangg gggaggtgtg      60
tcccaacatg anggtgnngt tctcttttga angaggggttg ngtttttann ccnggtgggt      120
gattnaaccc cattgtatgg agnnaaagggn tttnagggat ttttcggctc ttatcagtat      180
ntanattcct gtnaatcgga aaatnatntt tcnnccggaa aatnttgctc ccatccgnaa      240
attnctcccg ggtagtgcac nttngggggg cngccangtt tcccaggctg ctanaatcgt      300
actaaagntt naagtgggan tncaaataaa aacctnncac agagnatccn tacccgactg      360
tnnnntnctt tcgccctntg actctgcnnng agcccaatac ccnngngnat gtcncccnng      420
nnngcgcncn tgaaannnnc tcgnggetnn gancatcang gggtttcgca tcaaaagcnn      480
cgtttcncat naaggcactt tngcctcacc caaccnctng ccctcnccca tttngccgctc      540
nggttcncct acgctnntng cncctnnntn ganattttnc ccgcctnggg naancctcct      600
gnaatgggta gggnccttntc ttttnaccnn gnggtntact aatcnncctc acgctnctt      660
tctnaccccc cccctctttt caatcccanc ggcnaatggg gtctccccnn cgangggggg      720
nnnccannnc c

```

```

<210> 29
<211> 822
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(822)
<223> n = A,T,C or G

```

```

<400> 29
actagtccag tgtggtggaa ttccattgtg ttggggncnc ttctatgant antnttagat      60
cgctcanacc tcacancctc ccnacnangc ctataangaa nannaataga nctgtncnnt      120

```



```

catgtaccag ggctattaga agcaagaagg aaggagggag ggcagagcgc cctgctgagc 120
aaciaaggac tcctgcagcc ttctctgtct gtctcttggc gcaggcacat ggggaggcct 180
cccgagggt gggggccacc agtccagggg tgggagcact acanggggtg ggagtgggtg 240
gtggctggtg cnaatggcct gncacanat cctacgattc ttgacacctg gatttcacca 300
ggggaccttc tgttctccca nggnaacttc nttnatctcn aaagaacaca actgtttctt 360
cngcanttct ggctgttcat ggaaagcaca ggtgtccnat ttnggctggg acttgggtaca 420
tatggttccg gccacactct ccntcnaa aagtaattca ccccccccn ccntctnttg 480
cctggggcct taantacca caccggaact canttantta ttcattctng gntgggcttg 540
ntnatcnccn cctgaangcg ccaagttgaa aggccacgcc gtncnctc cccatagnan 600
nttttnnct canctaagc cccccnggc aacnatccaa tcccccccn tgggggcccc 660
agcccanggc ccccgncctg ggnnnccngn cncgnantcc ccaggntctc ccantcngnc 720
ccnnngcncc cccgcacgca gaacanaagg ntngagccnc cgcannnnnn nggttnncac 780
ctgcccccc ccnncgngg

```

<210> 32

<211> 789

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(789)

<223> n = A,T,C or G

<400> 32

```

tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 60
ttttnccnag ggcaggttta ttgacaacct cncgggacac aancaggctg gggacaggac 120
ggcaacaggc tccggcgcg gcggcggcg ccctacctgc ggtaccaa atngcagcctc 180
cgctcccgt tgatnttct ctgcagctgc aggatgcctt aaaacagggc ctcgccntn 240
ggtgggcacc ctgggatttn aatttccacg ggcacaatgc ggtcgcancc cctcaccacc 300
nattaggaat agtggnttta ccnccnccg ttggcnact cccntggaa accacttntc 360
gcggtccgg catctggtct taaaccttgc aaacnctggg gccctctttt tggttantnt 420
nccngccaca atcatnactc agactggcnc gggctggccc caaaaaancn ccccaaaacc 480
ggncatgtc ttncgggggt tgcctgnatn tncatcacct cccgggcnca ncaggncaac 540
ccaaaagttc ttngggcccn caaaaaanct cgggggggnc ccagtttcaa caaagtcac 600
ccccttggcc cccaaatcct cccccgntt nctgggtttg ggaaccacg cctctnnctt 660
tggnnngcaa gntggntccc ccttcgggcc ccgggtgggc ccnctctaa ngaaaacncc 720
ntcctnnnca ccatcccccc nngnnaacgnc tancaangna tccctttttt tanaaacggg 780
ccccccnccg

```

<210> 33

<211> 793

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(793)

<223> n = A,T,C or G

<400> 33


```

ggggatctct anatchnacct gnatgcatgg ttgtcgggtgt ggtecgctgtc gatgaanatg      60
aacaggatct tgccttgaa gctctcggt gctgtnttta agttgctcag tctgccgtca      120
tagtcagaca cncctctggg caaaaaacan caggatntga gtcttgattt cacctccaat      180
aatcttcngg gctgtctgct cggatgaactc gatgacnang ggcagctggg tgtgtntgat      240
aaantccanc angttctcct tggatgaactc cctttcaaag ttgttcgggc cttcatcaaa      300
cttctnnaan angannancc canctttgtc gagctggnat ttgganaaca cgtcactgtt      360
ggaaactgat cccaaatggg atgtcatcca tgcctctgtc tgcctgcaaa aaacttgctt      420
ggcncaaadc cgactcccn tcttgaaag aagccnatca cccccctc cctggactcc      480
nncaangact ctncgctnc cccntccnng cagggttggg ggcannccgg gcccttgctc      540
ttcttcagcc agttcacnat ntcatcagc cctctgcca gctgtntat tcttggggg      600
ggaanccgtc tctcccttcc tgaannaact ttgacgtng gaatagccgc gcntcncnt      660
acntnctggg ccgggttcaa antccctcnc ttgncnntcn cctcgggcca ttctggattt      720
nccnaacttt tctcttcccc cncctcngg ngtttgntt tttcatnggg ccccaactct      780
gctnttggcc antccctgg gggcntntan cncctcctnt ggtccctng ggcc      834

```

<210> 36

<211> 814

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(814)

<223> n = A,T,C or G

<400> 36

```

cggncgcttt ccngccgcgc ccggtttcca tgacnaaggc tcccttcang ttaaatacnn      60
cctagnaaac attaatgggt tgctctacta atacatcata cnaaccagta agcctgcca      120
naacgccaac tcaggccatt cctaccaaag gaagaaaggc tggctctctc acccctgta      180
ggaaaggcct gccttgtaag acaccacaat nccgctgaat ctnaagtctt gtgttttact      240
aatggaaaaa aaaaataaac aanagggtttt gttctcatgg ctgccaccg cagcctggca      300
ctaaaacanc ccagcgctca cttctgcttg ganaaatatt ctttgctctt ttggacatca      360
ggcttgatgg tatcactgcc acntttccac ccagctgggc ncccttcccc catntttgtc      420
antganctgg aaggcctgaa ncttagtctc caaaagtctc ngcccacaag accggccacc      480
aggggangtc ntttncagtg gatctgcaa anantaccn tatcatcnnt gaataaaaag      540
gcccctgaac ganatgcttc cancanctt taagacccat aatcctngaa ccatggtgcc      600
cttccggtct gatccnaaag gaatgttctt ggggtccant cctcctttg ttntttacgt      660
tgtnttggac cntgctngn atnaccnaan tganatcccc ngaagcacc tnccttggc      720
atttganttt cntaaattct ctgccctacn nctgaaagca cnattccctn ggcnccnaan      780
ggngaactca agaaggtctn ngaaaaacca cncn      814

```

<210> 37

<211> 760

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(760)

<223> n = A,T,C or G

<400> 37

```

gcatgctgct cttcctcaaa gttgttcttg ttgccataac aaccaccata ggtaaagcgg      60
gcgcagtggt cgctgaaggg gttgtagtac cagcgcggga tgctctcctt gcagagtcct    120
gtgtctggca ggtccacgca atgccctttg tctctgggga aatggatgcg ctggagctcg    180
tcnaanccac tcgtgtatth ttacangca gcctctccg aagcntccg gcagttgggg    240
gtgtcgtcac actccactaa actgtcgatn cancagccca ttgctgcagc ggaactgggt    300
gggctgacag gtgccagaac aactggatn ggcctttcca tggaagggcc tgggggaaat    360
cncctnancc caaactgcct ctcaaaggcc accttgaca ccccgacagg ctagaaatgc    420
actcttcttc ccaaaggtag ttgttcttgg tgcceaagca ncctccanca aacaaaaanc    480
ttgcaaaatc tgctccgtgg gggcatnnn taccanggtt ggggaaanaa acccggcngn    540
gancncctt gtttgaatgc naaggnaata atctctctgt cttgcttggg tggaanagca    600
caattgaact gttaacnttg ggccnggttc cncctnggtg gtctgaaact aatcacgcgc    660
actggaaaaa ggtangtgcc ttcttgaat tcccaaannt cccctngntt tgggtntttt    720
ctcctctncc ctaaaaatcg tnttcccccc cntanggcg      760

```

<210> 38

<211> 724

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(724)

<223> n = A,T,C or G

<400> 38

```

tttttttttt tttttttttt tttttttttt ttttttaaaa cccctccat tgaatgaaaa      60
cttcnaaat tgtccaaccc cctcnccaa atnnccattt ccgggggggg gttccaaacc    120
caaattaatt ttgganttta aattaaatnt tnattngggg aanaanccaa atgtnaagaa    180
aatttaaccc attatnaact taaatnccn gaaaccctg gnttccaaaa atttttaacc    240
cttaaatccc tccgaaattg ntaanggaaa accaaattcn cctaaggctn tttgaagggt    300
ngatttaaac ccccttnant tnttttnacc cngnctnaa ntatttngnt tccggtgttt    360
tcctnttaan cntnggtaac tcccgntaat gaannnccct aanccaatta aaccgaattt    420
tttttgaatt ggaaattccn ngggaattna ccgggggttt tccnttttg gggccatncc    480
ccnctttcg gggtttggn ntaggttgaa tttttnnang nccccaaaaa ncccccaana    540
aaaaaactcc caagnnttaa ttngaantnc ccccttccca ggccttttg gaaaggnggg    600
ttnttggggg ccngggantc cnttcccccn ttncncccc cccccnggt aaanggttat    660
ngnntttggt ttttgggccc cttnanggac cttccgcatn gaaattaaat ccccggnccg    720
gccg      724

```

<210> 39

<211> 751

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(751)

<223> n = A,T,C or G

<400> 39

<210> 45

acaagggggc	ataatgaagg	agtggggana	gattttaaag	aaggaaaaaa	aacgaggccc	60
tgaacagaat	tttctgnac	aacggggctt	caaaataatt	ttcttgggga	ggttcaagac	120
gcttcactgc	ttgaaactta	aatggatgtg	ggacanaatt	ttctgtaatg	accctgaggg	180
cattacagac	gggactctgg	gaggaaggat	aaacagaaaag	gggacaaagg	ctaataccaa	240
aacatcaaag	aaaggaaggt	ggcgtcatac	ctcccagcct	acacagttct	ccagggctct	300
cctcatccct	ggaggacgac	agtggaggaa	caactgacca	tgtccccagg	ctcctgtgtg	360
ctggctcctg	gtcttcagcc	cccagctctg	gaagcccacc	ctctgctgat	cctgcgtggc	420
ccacactcct	tgaacacaca	tccccaggtt	atattcctgg	acatggctga	acctcctatt	480

```

cctacttccg agatgccttg ctccctgcag cctgtcaaaa tcccactcac cctccaaacc 540
acggcatggg aagcctttct gacttgccctg attactccag catcttggaa caatccctga 600
ttccccactc cttagaggca agataggggtg gttaagagta gggctggacc acttgagacc 660
aggctgctgg cttcaaattn tggctcattt acgagctatg ggaccttggg caagtnatct 720
tcacttctat gggcntcatt ttgttctacc tgcaaaatgg gggataataa tagt 774

```

```

<210> 48
<211> 124
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(124)
<223> n = A,T,C or G

```

```

<400> 48
canaaattga aattttataa aaaggcattt ttctcttata tccataaaat gatataattt 60
ttgcaantat anaaatgtgt cataaattat aatgttcctt aattacagct caacgcaact 120
tggt 124

```

```

<210> 49
<211> 147
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(147)
<223> n = A,T,C or G

```

```

<400> 49
gccgatgcta ctatttttatt gcaggagggtg ggggtgtttt tattattctc tcaacagctt 60
tgtggctaca ggtgggtgtct gactgcatna aaaanttttt tacgggtgat tgcaaaaatt 120
ttagggcacc catatcccaa gcantgt 147

```

```

<210> 50
<211> 107
<212> DNA
<213> Homo sapien

```

```

<400> 50
acattaaatt aataaaaagga ctgttgggggt tctgctaaaa cacatggctt gatatatattgc 60
atggttttgag gttaggagga gttaggcata tgttttggga gaggggt 107

```

```

<210> 51
<211> 204
<212> DNA
<213> Homo sapien

```

```

<400> 51

```

```
<210> 52
<211> 491
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(491)
<223> n = A,T,C or G
```

```
<210> 53
<211> 484
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(484)
<223> n = A,T,C or G
```

```
<210> 54
<211> 151
<212> DNA
<213> Homo sapien
```


<400> 54
 actaaacctc gtgcttgtga actccataca gaaaacggtg ccatccctga acacggctgg 60
 ccactgggta tactgctgac aaccgcaaca acaaaaacac aaatccttgg cactggctag 120
 tctatgtcct ctcaagtgcc tttttgtttg t 151

<210> 55
 <211> 91
 <212> DNA
 <213> Homo sapien

<400> 55
 acctggcttg tctccgggtg gttcccggtg cccccacgg tccccagaac ggacactttc 60
 gccctccagt ggatactcga gccaaagtgg t 91

<210> 56
 <211> 133
 <212> DNA
 <213> Homo sapien

<400> 56
 ggcggatgtg cgttggttat atacaaatat gtcattttat gtaagggact tgagtatact 60
 tggatttttg gtatctgtgg gttgggggga cggctccagga accaataccc catggatacc 120
 aagggacaac tgt 133

<210> 57
 <211> 147
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(147)
 <223> n = A,T,C or G

<400> 57
 actctggaga acctgagccg ctgctccgcc tctgggatga ggtgatgcan gcngtggcgc 60
 gactgggagc tgagcccttc cctttgcgcc tgcctcagag gattgttgcc gacntgcana 120
 tctcantggg ctggatncat gcagggt 147

<210> 58
 <211> 198
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(198)
 <223> n = A,T,C or G

<400> 58

```

acagggatat aggtttnaag ttattgtnat tgtaaaatac attgaatttt ctgtatactc      60
tgattacata catttatcct ttaaaaaaga tgtaaatctt aatttttatg ccatctatta      120
atttaccaat gagttacctt gtaaatgaga agtcatgata gcactgaatt ttaactagtt      180
ttgacttcta agtttggt                                     198

```

```

<210> 59
<211> 330
<212> DNA
<213> Homo sapien

```

```

<400> 59
acaacaaatg ggttgtgagg aagtcttatac agcaaaactg gtgatggcta ctgaaaagat      60
ccattgaaaa ttatcattaa tgatttttaa tgacaagtta tcaaaaactc actcaatttt      120
cacctgtgct agcttgctaa aatgggagtt aactctagag caaatatagt atcttctgaa      180
tacagtcaat aaatgacaaa gccagggcct acaggtgggt tccagacttt ccagaccag      240
cagaaggaat ctattttatac acatggatct ccgtctgtgc tcaaaatacc taatgatatt      300
tttcgtcttt attggacttc tttgaagagt                                     330

```

```

<210> 60
<211> 175
<212> DNA
<213> Homo sapien

```

```

<400> 60
accgtgggtg ccttctacat tcctgacggc tccttcacca acatctgggt ctacttcggc      60
gtcgtgggct ccttctctt catctcctc cagctgggtgc tgctcatcga ctttgcgcac      120
tcctggaacc agcgggtggct gggcaaggcc gaggagtgcg attcccgtgc ctgggt      175

```

```

<210> 61
<211> 154
<212> DNA
<213> Homo sapien

```

```

<400> 61
acccactttt tcctcctgtg agcagtctgg acttctcact gctacatgat gagggtgagt      60
ggttgtgtgt cttcaacagt atcctccctt ttccggatct gctgagccgg acagcagtgc      120
tggaactgcac agccccgggg ctccacattg ctgt                                     154

```

```

<210> 62
<211> 30
<212> DNA
<213> Homo sapien

```

```

<400> 62
cgctcgagcc ctatagttag tcgtattaga                                     30

```

```

<210> 63
<211> 89
<212> DNA
<213> Homo sapien

```



```

actacacaca ctccacttgc ccttgtgaga cactttgtcc cagcacttta ggaatgctga      60
ggtcggacca gccacatctc atgtgcaaga ttgccagca gacatcaggt ctgagagttc      120
cccttttaaa aaaggggact tgcttaaaaa agaagtctag ccacgattgt gtagagcagc      180
tgtgctgtgc tggagattca cttttgagag agttctcctc tgagacctga tcttttagagg      240
ctgggcagtc ttgcacatga gatggggctg gtctgatctc agcactcctt agtctgcttg      300
cctctcccag ggccccagcc tggccacacc tgcttacagg gcactctcag atgcccatac      360
catagtttct gtgctagtgg accgt                                           385

```

```

<210> 68
<211> 73
<212> DNA
<213> Homo sapien

```

```

<400> 68
acttaaccag atatattttt accccagatg gggatattct ttgtaaaaaa tgaaaataaa      60
gttttttttaa tgg                                                         73

```

```

<210> 69
<211> 536
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(536)
<223> n = A,T,C or G

```

```

<400> 69
actagtcag tgtggtggaa ttccattgtg ttgggggctc tcaccctcct ctctgcagc      60
tccagctttg tgctctgcct ctgaggagac catggcccag catctgagta ccctgctgct      120
cctgctggcc accctagctg tggccctggc ctggagcccc aaggaggagg ataggataat      180
cccggtggc atctataacg cagacctcaa tgatgagtgg gtacagcgtg cccttcactt      240
cgccatcagc gagtataaca aggccaccaa agatgactac tacagacgtc cgctgcgggt      300
actaagagcc aggcaacaga ccgttggggg ggtgaattac ttcttcgacg tagaggtggg      360
ccgaaccata tgtaccaagt ccagcccaa cttggacacc tgtgccttcc atgaacagcc      420
agaactgcag aagaaacagt tgtgctcttt cgagatctac gaagttccct ggggagaaca      480
gaangtcctt gggtgaaatc caggtgtcaa gaaatcctan ggatctgttg ccaggc      536

```

```

<210> 70
<211> 477
<212> DNA
<213> Homo sapien

```

```

<400> 70
atgacccta acaggggccc tctcagccct cctaattgacc tccggcctag ccatgtgatt      60
tacttccac tccataacgc tctcataact aggcctacta accaaccacac taaccatata      120
ccaatgatgg cgcgatgtaa cagagaaag cacataccaa ggccaccaca caccacctgt      180
ccaaaaaggc cttcgatacg ggataatcct atttattacc tcagaagttt ttttcttcgc      240
agggttttt ctgagccttt taccactcca gcctagcccc taccocccaa ctaggagggc      300
actggcccc aacaggcatc acccgcgtaa atcccctaga agtcccactc ctaaaccacat      360
ccgtattact cgcacagga gtatcaatca cctgagctca ccatagtcta atagaaaaca      420
accgaaacca aattattcaa agcactgctt attacaattt tactgggtct ctatttt      477

```


<223> n = A,T,C or G

<400> 73

cagtgccagc	actggtgcc	gtaccagtac	caataacagt	gccagtgcc	gtgccagcac	60
cagtgggtggc	ttcagtgtctg	gtgccagcct	gaccgccact	ctcacatttg	ggctcttcgc	120
tggccttggt	ggagctggtg	ccagcaccag	tggcagctct	ggcgcctgtg	gtttctccta	180
caagtgagat	tttagatatt	gttaatcctg	ccagtctttc	tcttcaagcc	aggggtgcatc	240
ctcagaaacc	tactcaacac	agcactctag	gcagccacta	tcaatcaatt	gaagttgaca	300
ctctgcatta	aatctatttg	ccatttctga	aaaaaaaaaa	aaaaaaagg	cggccgctcg	360
antctagagg	gccccgttaa	acccgctgat	cagcctcgac	tgtgccttct	anttgccagc	420
catctgttgt	ttgccccctc	cccngtgcct	tccttgaccc	tggaaagtgc	cactccctc	480
gtccttttct	aantaaaat					499

<210> 74

<211> 537

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(537)

<223> n = A,T,C or G

<400> 74

tttcatagga	gaacacactg	aggagatact	tgaagaatth	ggattcagcc	gcgaagagat	60
ttatcagctt	aactcagata	aatcattga	aagtaataag	gtaaaagcta	gtctctaact	120
tccaggccca	cggctcaagt	gaatttgaat	actgcattta	cagtgtagag	taacacataa	180
cattgtatgc	atggaaacat	ggaggaacag	tattacagtg	tcctaccact	ctaatacaaga	240
aaagaattac	agactctgat	tctacagtga	tgattgaatt	ctaaaaatgg	taatcattag	300
ggcttttgat	ttataanact	ttgggtactt	atactaaatt	atggtagtta	tactgccttc	360
cagtttgctt	gatatatttg	ttgatattaa	gattcttgac	ttatatattg	aatgggttct	420
actgaaaaan	gaatgatata	ttcttgaaga	catcgatata	catttattta	cactcttgat	480
tctacaatgt	agaaaatgaa	ggaaatgccc	caaattgtat	ggtgataaaa	gtccccgt	537

<210> 75

<211> 467

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(467)

<223> n = A,T,C or G

<400> 75

caaanacaat	tggttcaaaag	atgcaaata	tacactactg	ctgcagctca	caaacacctc	60
tgcataattac	acgtacctcc	tcctgctcct	caagtagtgt	ggctctatttt	gccatcatca	120
cctgctgtct	gcttagaaga	acggctttct	gctgcaangg	agagaaatca	taacagacgg	180
tggcacaagg	aggccatctt	ttcctcatcg	gttattgtcc	ctagaagcgt	cttctgagga	240
tctagtggg	ctttctttct	gggtttgggc	catttcantt	ctcatgtgtg	tactattcta	300
tcattattgt	ataacgggtt	tcaaacngt	gggcacncag	agaacctcac	tctgtaataa	360


```

<221> misc_feature
<222> (1)...(552)
<223> n = A,T,C or G

```

```

<400> 79
tccttttgtt aggtttttga gacaacccta gacctaaact gtgtcacaga cttctgaatg      60
tttaggcagt gctagtaatt tcctcgtaat gattctgtta ttactttcct attctttatt      120
cctctttcct ctgaagatta atgaagtga aaattgaggt ggataaatac aaaaaggtag      180
tgtgatagta taagtatcta agtgcagatg aaagtgtgtt atatatatcc attcaaaatt      240
atgcaagtta gtaattactc aggggttaact aaattacttt aatatgctgt tgaacctact      300
ctgttccttg gctagaaaaa attataaaca ggactttgtt agtttgggaa gccaaattga      360
taatattcta tgttctaaaa gttgggctat acataaanta tnaagaaata tggaatttta      420
ttcccaggaa tatgggggttc atttatgaat antaccggg anagaagttt tgantnaaac      480
cngtttttgt taatacgta atatgtcctn aatnaacaag gcntgactta tttccaaaaa      540
aaaaaaaaaa aa                                     552

```

```

<210> 80
<211> 476
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(476)
<223> n = A,T,C or G

```

```

<400> 80
acagggattt gagatgctaa ggccccagag atcgtttgat ccaaccctct tattttcaga      60
ggggaaaatg gggcctagaa gttacagagc atctagctgg tgcgctggca cccctggcct      120
cacacagact cccgagtagc tgggactaca ggcacacagt cactgaagca ggccctgttt      180
gcaattcacg ttgccacctc caacttaaac attcttcata tgtgatgtcc ttagtacta      240
aggttaaact ttcccacca gaaaaggcaa cttagataaa atcttagagt actttcatac      300
tcttctaagt cctcttcag cctcactttg agtcctcctt ggggggtgat aggaantntc      360
tcttggtttt ctcaataaaa tctctatcca tctcatgttt aatttggtac gcntaaaaat      420
gctgaaaaaa ttaaaatggt ctggtttcnc tttaaaaaaa aaaaaaaaaa aaaaaa      476

```

```

<210> 81
<211> 232
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(232)
<223> n = A,T,C or G

```

```

<400> 81
tttttttttg tatgccntcn ctgtgnggtt attgttgctg ccaccctgga ggagcccagt      60
ttcttctgta tctttctttt ctggggggtc ttcttggtc tgcccctcca ttcccagcct      120
ctcatcccca tcttgcaact ttgctagggg tggaggcgct ttcttggtag cccctcagag      180
actcagtcag cgggaataag tcctaggggt ggggggtgtg gcaagccggc ct                232

```


<210> 82
 <211> 383
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(383)
 <223> n = A,T,C or G

<400> 82
 aggcgggagc agaagctaaa gccaaagccc aagaagagtg gcagtgccag cactggtgcc 60
 agtaccagta ccaataacat gccagtgcc gtgccagcac cagtgggtggc ttcagtgtg 120
 gtgccagcct gaccgccact ctcacatttg ggctcttcgc tggccttggg ggagctggtg 180
 ccagcaccag tggcagctct ggtgcctgtg gtttctccta caagtgagat tttagatatt 240
 gttaatcctg ccagtctttc ttttcaagcc aggggtgcac ctcagaaacc tactcaacac 300
 agcactctng gcagccacta tcaatcaatt gaagttgaca ctctgcatta aatctatttg 360
 ccatttcaaa aaaaaaaaaa aaa 383

<210> 83
 <211> 494
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(494)
 <223> n = A,T,C or G

<400> 83
 accgaattgg gaccgctggc ttataagcga tcatgtcttc cagtattacc tcaacgagca 60
 gggagatcga gtctatacgc tgaagaaatt tgaccgatg ggacaacaga cctgctcagc 120
 ccatcctgct cggttctccc cagatgacaa atactctcga caccgaatca ccatcaagaa 180
 acgcttcaag gtgctcatga cccagcaacc gcgcctgtc ctctgagggg ccttaaactg 240
 atgtcttttc tgccacctgt taccctcgg agactccgta accaaactct tcggactgtg 300
 agccctgatg cctttttgcc agccatactc tttggcntcc agtctctcgt ggcgattgat 360
 tatgcttgty tgaggcaatc atgggtggcat caccatnaa gggaacacat ttganttttt 420
 tttcncatat tttaaattac naccagaata nttcagaata aatgaattga aaaactctta 480
 aaaaaaaaaa aaaa 494

<210> 84
 <211> 380
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(380)
 <223> n = A,T,C or G

<400> 84

```

gctggtagcc tatggcgtgg ccacggangg gctcctgagg cacgggacag tgacttccca      60
agtatcctgc gccgcgtctt ctaccgtccc tacctgcaga tcttcgggca gattccccag      120
gaggacatgg acgtggccct catggagcac agcaactgct cgtcggagcc cggcttcttg      180
gcacaccctc ctggggccca ggccggcacc tgcgtctccc agtatgcaa ctggctggtg      240
gtgctgctcc tcgtcatctt cctgctcgtg gccaacatcc tgctggtcac ttgctcattg      300
ccatgttcag ttacacattc ggcaaagtac agggcaacag cnatctctac tgggaaggcc      360
agcgttnccg cctcatccgg

```

<210> 85

<211> 481

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(481)

<223> n = A,T,C or G

<400> 85

```

gagttagctc ctccacaacc ttgatgaggt cgtctgcagt ggctctctgc ttcataccgc      60
tnccatcgtc atactgtagg ttggccacca cctcctgcat cttggggcgg ctaatatcca      120
ggaaactctc aatcaagtca ccgtcnatna aacctgtggc tggttctgtc ttccgctcgg      180
tgtgaaagga tctccagaag gagtgtctga tcttccccac acttttgatg actttattga      240
gtcgattctg catgtccagc aggaggttgt accagctctc tgacagtgag gtcaccagcc      300
ctatcatgcc nttgaacgtg ccgaagaaca ccgagccttg tgtggggggg gnagtctcac      360
ccagattctg cattaccaga nagccgtggc aaaaganatt gacaactcgc ccaggngaa      420
aaagaacacc tcctggaagt gctngccgct cctcgtccnt tggtggnngc gcntnccttt      480
t

```

<210> 86

<211> 472

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(472)

<223> n = A,T,C or G

<400> 86

```

aacatcttcc tgtataatgc tgtgtaatat cgatccgatn ttgtctgctg agaattcatt      60
acttgaaaaa gcaacttnaa gcctggacac tgggtattaaa attcacaata tgcaaacatt      120
taaacagtgt gtcaatctgc tcccttactt tgtcatcacc agtctgggaa taagggtatg      180
ccctattcac acctgttaaa agggcgctaa gcatttttga ttcaacatct ttttttttga      240
cacaagtccg aaaaaagcaa aagtaaacag ttnttaattt gttagccaat tcactttctt      300
catgggacag agccatttga tttaaaaagc aaattgcata atattgagct ttgggagctg      360
atatntgagc ggaagantag cctttctact tcaccagaca caactccttt catattggga      420
tgttnacnaa agttatgtct cttacagatg ggatgctttt gtggcaattc tg      472

```

<210> 87

<211> 413
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(413)
 <223> n = A,T,C or G

<400> 87
 agaaaccagt atctctnaaa acaacctctc ataccttgtg gacctaatth tgtgtgctg 60
 tgtgtgtgcg cgcataattat atagacaggg acatcttttt tacttttgta aaagcttatg 120
 cctcttttgt atctatatct gtgaaagttt taatgatctg ccataatgtc ttggggacct 180
 ttgtcttctg tgtaaatggg actagagaaa acacctatnt tatgagtcaa tctagttngt 240
 tttattcgac atgaaggaaa tttccagatn acaacactna caaactctcc cttgactagg 300
 ggggacaaaag aaaagcnaaa ctgaacatna gaaacaattn cctgggtgaga aattncataa 360
 acagaaattg ggtngtatat tgaaanannn catcattnaa acgttttttt ttt 413

<210> 88
 <211> 448
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(448)
 <223> n = A,T,C or G

<400> 88
 cgcagcgggt cctctctatc tagctccage ctctcgctg cccactccc cgcgtcccgc 60
 gtcctagccn accatggccg ggccccctgc cgccccgctg ctctgctgg ccctcctggc 120
 cgtggccctg gccgtgagcc ccgcggccgg ctccagtcce ggcaagccgc cgcgcctggg 180
 gggaggccca tggacccccg gtggaagaag aagggtgtgc gcgtgcactg gactttgccc 240
 tcggcnanta caacaaaccc gcaacnactt ttaccnagcn cgcgctgcag gttgtgccc 300
 cccaancaa ttgttactng gggtaantaa ttcttggaag ttgaacctgg gccaaacnng 360
 tttaccagaa ccnagccaat tngaacaatt nccccccat aacagcccct tttaaaaagg 420
 gaancantcc tgntcttttc caaattht 448

<210> 89
 <211> 463
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(463)
 <223> n = A,T,C or G

<400> 89
 gaattttgtg cactggccac tgtgatggaa ccattgggce aggatgcttt gagtttatca 60
 gtagtgattc tgccaaagtt ggtgtgtgtaa catgagtatg taaaatgtca aaaaattagc 120

<220>
 <221> misc_feature
 <222> (1)...(477)
 <223> n = A,T,C or G

<400> 92
 atacagccca natcccacca cgaagatgcg cttgttgact gagaacctga tgcggtcact 60
 ggtcccgcgtg tagccccagc gactctccac ctgctggaag cggttgatgc tgcactcctt 120
 cccacgcagg cagcagcggg gccggtcaat gaactccact cgtggcttgg ggttgacggg 180
 taantgcagg aagaggctga ccacctcgcg gtccaccagg atgcccgact gtgcgggacc 240
 tgcagcgaaa ctctcgatg gtcatgagcg ggaagcgaat gangcccagg gccttgccca 300
 gaaccttccg cctgttctct ggcgtcacct gcagctgctg ccgctnacac tcggcctcgg 360
 accagcggac aaacggcggt gaacagccgc acctcacgga tgcccantgt gtcgcgctcc 420
 aggaacggcn ccagcgtgtc caggtcaatg tcggtgaanc ctccgcgggt aatggcg 477

<210> 93
 <211> 377
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(377)
 <223> n = A,T,C or G

<400> 93
 gaacggctgg accttgctc gcattgtgct gctggcagga ataccttggc aagcagctcc 60
 agtccgagca gcccagacc gctgccgcc gaagctaagc ctgcctctgg cttccccctc 120
 cgcctcaatg cagaaccant agtgggagca ctgtgttttag agttaagagt gaacactgtn 180
 tgattttact tgggaatttc ctctgttata tagcttttcc caatgctaata ttccaaacaa 240
 caacaacaaa ataacatgtt tgctgtttna gttgtataaa agtangtgat tctgtatnta 300
 aagaaaatat tactgttaca tatactgctt gcaanttctg tatttattgg tncctctggaa 360
 ataaatatat tattaata 377

<210> 94
 <211> 495
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(495)
 <223> n = A,T,C or G

<400> 94
 ccctttgagg ggttagggtc cagttcccag tggaagaaac aggccaggag aantgcgtgc 60
 cgagctgang cagatttccc acagtgacc cagagccctg ggctatagtc tctgaccctt 120
 ccaaggaaaag accaccttct ggggacatgg gctggagggc aggacctaga ggcaccaagg 180
 gaaggcccca ttccggggct gttccccgag gaggaaggga aggggctctg tgtgcccccc 240
 acgaggaana ggccctgant cctgggatca nacaccctt cacgtgtatc cccacacaaa 300
 tgcaagctca ccaaggtccc ctctcagtc cttccctaca ccctgaacgg nactggccc 360

```
<220>
<221> misc feature
```

<222> (1)...(479)

<223> n = A,T,C or G

<400> 97

actcttttcta	atgctgatata	gatcttgagt	ataagaatgc	atatgtcact	agaatggata	60
aaataatgct	gcaaacttaa	tgttcttatg	caaaatggaa	cgctaataa	acacagctta	120
caatcgcaaa	tcaaaaactca	caagtgtcca	tctgtttag	atttagtgta	ataagactta	180
gattgtgctc	cttcggatat	gattgtttct	canatcttgg	gcaatnttcc	ttagtcaaata	240
caggctacta	gaattctgtt	attggatatn	tgagagcatg	aaatttttaa	naatacactt	300
gtgattatna	aattaatcac	aaatttcact	tatacctgct	atcagcagct	agaaaaacat	360
ntnnttttta	natcaaagta	ttttgtgttt	ggaantgttn	aaatgaaatc	tgaatgtggg	420
ttcnatctta	ttttttcccn	gacnactant	tnctttttta	gggnctattc	tgancctatc	479

<210> 98

<211> 461

<212> DNA

<213> Homo sapien

<400> 98

agtgacttgt	cctccaacaa	aacccttga	tcaagtttgt	ggcactgaca	atcagacctta	60
tgctagtcc	tgctatctat	tcgtactaa	atgcagactg	gaggggacca	aaaaggggca	120
tcaactccag	ctggattatt	ttggagcctg	caaactctatt	cctacttgta	cggactttga	180
agtgattcag	tttctctac	ggatgagaga	ctggctcaag	aatatcctca	tgacgcttta	240
tgaagccact	ctgaacacgc	tggttatcta	gatgagaaca	gagaaataaa	gtcagaaaat	300
ttacctggag	aaaagaggct	ttggctgggg	accatcccat	tgaaccttct	cttaaggact	360
ttaagaaaaa	ctaccacatg	ttgtgtatcc	tggtgccggc	cgtttatgaa	ctgaccaccc	420
tttgaataaa	tcttgacgct	cctgaacttg	ctcctctgcg	a		461

<210> 99

<211> 171

<212> DNA

<213> Homo sapien

<400> 99

gtggccgcgc	gcaggtgttt	cctcgtaacc	cagggccccc	tccttcccc	aggcgtccct	60
cggcgctct	gcgggcccga	ggaggagcgg	ctggcggtg	gggggagtgt	gaccaccct	120
cggtgagaaa	agccttctct	agcgatctga	gaggcggtgc	ttgggggtac	c	171

<210> 100

<211> 269

<212> DNA

<213> Homo sapien

<400> 100

cggccgcaag	tgcaactcca	gctggggccg	tgccgacgaa	gattctgcca	gcagttggtc	60
cgactgcgac	gacggcgggc	gcgacagtcg	caggtgcagc	gcggggcgct	ggggtcttgc	120
aaggctgagc	tgacgccgca	gaggtcgtgt	cacgtcccac	gaccttgacg	ccgtcgggga	180
cagccggaac	agagcccggg	gaagcgggag	gcctcgggga	gcccctcggg	aaggcgggcc	240
cgagagatac	gcaggtgcag	gtggccgcgc				269

<210> 101

<211> 405
 <212> DNA
 <213> Homo sapien

<400> 101
 tttttttttt ttttggaaac tactgcgagc acagcaggtc agcaacaagt ttattttgca 60
 gctagcaagg taacagggtg gggcatgggt acatgttcag gtcaacttcc tttgtcgtgg 120
 ttgattgggt tgtctttatg ggggcggggg ggggtagggg aaacgaagca aataacatgg 180
 agtgggtgca ccctccctgt agaacctggg tacaagctt ggggcagttc acctggtctg 240
 tgaccgtcat tttcttgaca tcaatgttat tagaagtcag gatatctttt agagagtcca 300
 ctgttctgga gggagattag ggtttcttgc caaatccaac aaaatccact gaaaaagttg 360
 gatgatcagt acgaataacc aggcatactt tcatatcggt ggcca 405

<210> 102
 <211> 470
 <212> DNA
 <213> Homo sapien

<400> 102
 tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 60
 ggcacttaat ccatttttat ttcaaaatgt ctacaaatth aatcccatta tacgggtattt 120
 tcaaaatcta aattattcaa attagccaaa tccttaccaa ataataccca aaaatcaaaa 180
 atatacttct ttcagcaaac ttgttacata aattaaaaaa atatatacgg ctggtgtttt 240
 caaagtacaa ttatcttaac actgcaaaca ttttaaggaa ctaaaataaa aaaaaacact 300
 ccgcaaaggt taaagggaac aacaaattct tttacaacac cattataaaa atcatatctc 360
 aaatcttagg ggaatatata cttcacacgg gatcttaact tttactcact ttgtttattt 420
 ttttaaacca ttgtttgggc ccaacacaat ggaatcccc ctggactagt 470

<210> 103
 <211> 581
 <212> DNA
 <213> Homo sapien

<400> 103
 tttttttttt ttttttttga cccccctctt ataaaaaaca agttaccatt ttattttact 60
 tacacatatt tattttataa ttgggtattag atattcaaaa ggcagctttt aaaatcaaac 120
 taaatggaaa ctgccttaga tacataattc ttaggaatta gcttaaaatc tgcctaaagt 180
 gaaaatcttc tctagctctt ttgactgtaa attttttgact cttgtaaaac atccaaattc 240
 atttttcttg tctttaaatt tatctaattc ttccattttt tccctattcc aagtcaattt 300
 gcttctctag cctcatttcc tagctcttat ctactattag taagtggctt ttttcctaaa 360
 agggaaaaca ggaagagaaa tggcacacaa aacaaacatt ttatattcat atttctacct 420
 acgttaataa aatagcattt tgtgaagcca gctcaaaaga aggcttagat ccttttatgt 480
 ccatttttagt cactaaacga tatcaaagtg ccagaatgca aaaggtttgt gaacatttat 540
 tcaaaagcta atataagata tttcacatac tcatctttct g 581

<210> 104
 <211> 578
 <212> DNA
 <213> Homo sapien

<400> 104

002290-002290

tttttttttt	tttttttttt	tttttctctt	cttttttttt	gaaatgagga	tcgagttttt	60
cactctctag	atagggcatg	aagaaaactc	atctttccag	ctttaaaata	acaatcaa	120
ctcttatgct	atatcatatt	ttaagttaaa	ctaattgagtc	actggcttat	cttctcctga	180
aggaaatctg	ttcattcttc	tcattcatat	agttatatca	agtactacct	tgcattattga	240
gaggtttttc	ttctctat	acacatatat	ttccatgtga	atttgtatca	aacctttatt	300
ttcatgcaaa	ctagaaaata	atgtttcttt	tgcataagag	aagagaacaa	tatagcatta	360
caaaactgct	caaattgttt	gttaagttaa	ccattataat	tagttggcag	gagctaatac	420
aaatcacatt	tacgacagca	ataataaaac	tgaagtacca	gttaaataatc	caaaataatt	480
aaaggaacat	ttttagcctg	ggtataatta	gctaattcac	tttacaagca	tttattagaa	540
tgaattcaca	tgttattatt	cctagcccaa	cacaatgg			578

<210> 105

<211> 538

<212> DNA

<213> Homo sapien

<400> 105

tttttttttt	tttttcagta	ataatcagaa	caatatttat	ttttatattt	aaaattcata	60
gaaaagtgcc	ttacatttaa	taaaagtttg	tttctcaaag	tgatcagagg	aattagatat	120
gtcttgaaca	ccaatattaa	tttgaggaaa	atacaccaaa	atacattaag	taaattattt	180
aagatcatag	agcttgtaag	tgaaaagata	aaatttgacc	tcagaaaactc	tgagcattaa	240
aaatccacta	ttagcaata	aattactatg	gacttcttgc	tttaattttg	tgatgaatat	300
ggggtgtcac	tggtaaacca	acacattctg	aaggatacat	tacttagtga	tagattctta	360
tgtactttgc	taatacgtgg	atatgagttg	acaagtttct	ctttcttcaa	tcttttaagg	420
ggcgagaaat	gaggaagaaa	agaaaaggat	tacgcatact	gttcttttcta	tggaaggatt	480
agatatgttt	cctttgccaa	tattaaaaaa	ataataatgt	ttactactag	tgaaaccc	538

<210> 106

<211> 473

<212> DNA

<213> Homo sapien

<400> 106

tttttttttt	tttttttagtc	aagtttctat	ttttattata	attaaagtct	tggtcatttc	60
atttatttagc	tctgcaactt	acatatattaa	attaaagaaa	cgtttttagac	aactgtacaa	120
tttataaatg	taaggtgcca	ttattgagta	atatattcct	ccaagagtgg	atgtgtccct	180
tctcccacca	actaatgaac	agcaacatta	gtttaatttt	attagtagat	atacactgct	240
gcaaacgcta	attctcttct	ccatcccat	gtgatattgt	gtatatgtgt	gagttggtag	300
aatgcatcac	aatctacaat	caacagcaag	atgaagctag	gctgggcttt	cggtgaaaat	360
agactgtgtc	tgtctgaatc	aaatgatctg	acctatcctc	ggtggcaaga	actcttcgaa	420
ccgcttcctc	aaaggcgctg	ccacatttgt	ggctctttgc	acttgtttca	aaa	473

<210> 107

<211> 1621

<212> DNA

<213> Homo sapien

<400> 107

cgccatggca	ctgcagggca	tctcggtcat	ggagctgtcc	ggcctggccc	cgggcccgtt	60
ctgtgtcatg	gtcctggctg	acttcggggc	gcgtgtggta	cgcgtggacc	ggcccggctc	120
ccgctacgac	gtgagccgct	tgggcccggg	caagcgctcg	ctagtgtctg	acctgaagca	180

```

gccgcgggga gccgccgtgc tgcggcggtct gtgcaagcgg tcggatgtgc tgctggagcc 240
cttccgccgc ggtgtcatgg agaaactcca gctgggcccc gagattctgc agcgggaaaa 300
tccaaggctt atttatgcca ggctgagtgg atttggccag tcaggaagct tctgccggtt 360
agctggccac gatatcaact atttggcttt gtcagggtgtt ctctcaaaaa ttggcagaag 420
tggtgagaat ccgtatgccc cgctgaatct cctggctgac tttgctggtg gtggccttat 480
gtgtgactg ggcattataa tggctctttt tgaccgcaca cgcactgaca agggtcaggt 540
cattgatgca aatatggtgg aaggaacagc atatttaagt tcttttctgt ggaaaaactca 600
gaaatcgagt ctgtgggaag cacctcgagg acagaacatg ttggatggtg gagcaccttt 660
ctatacgact tacaggacag cagatgggga attcatggct gttggagcaa tagaacccca 720
gttctacgag ctgctgatca aaggacttgg actaaagtct gatgaacttc ccaatcagat 780
gagcatggat gattggccag aaatgaagaa gaagtttgca gatgtatttg caaagaagac 840
gaaggcagag tgggtgtcaaa tctttgacgg cacagatgcc tgtgtgactc cggttctgac 900
ttttgaggag gttgttcac atgatcacia caaggaacgg ggctcgttta tcaccagtga 960
ggagcaggac gtgagcccc gccctgcacc tctgctgtta aacaccccag ccatcccttc 1020
tttcaaaagg gatcctttca taggagaaca cactgaggag atacttgaag aatttggatt 1080
cagccgcgaa gagatttatc agcttaactc agataaaatc attgaaagta ataaggtaaa 1140
agctagtctc taacttccag gccacgggt caagtgaatt tgaatactgc atttacagt 1200
tagagtaaca cataacattg tatgcatgga aacatggagg aacagtatta cagtgtccta 1260
ccactcta at caagaaaaga attacagact ctgattctac agtgatgatt gaattctaaa 1320
aatggttatc attagggctt ttgatttata aaactttggg tacttatact aaattatggt 1380
agttattctg ccttccagtt tgcttgatat atttgttgat attaagattc ttgacttata 1440
ttttgaatgg gttctagtga aaaaggaatg atatattctt gaagacatcg atatacat 1500
atttacactc ttgattctac aatgtagaaa atgaggaaat gccacaaatt gtatggtgat 1560
aaaagtcacg tgaaacaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1620
a 1621

```

<210> 108

<211> 382

<212> PRT

<213> Homo sapien

<400> 108

```

Met Ala Leu Gln Gly Ile Ser Val Met Glu Leu Ser Gly Leu Ala Pro
1          5          10          15
Gly Pro Phe Cys Ala Met Val Leu Ala Asp Phe Gly Ala Arg Val Val
20          25          30
Arg Val Asp Arg Pro Gly Ser Arg Tyr Asp Val Ser Arg Leu Gly Arg
35          40          45
Gly Lys Arg Ser Leu Val Leu Asp Leu Lys Gln Pro Arg Gly Ala Ala
50          55          60
Val Leu Arg Arg Leu Cys Lys Arg Ser Asp Val Leu Leu Glu Pro Phe
65          70          75          80
Arg Arg Gly Val Met Glu Lys Leu Gln Leu Gly Pro Glu Ile Leu Gln
85          90          95
Arg Glu Asn Pro Arg Leu Ile Tyr Ala Arg Leu Ser Gly Phe Gly Gln
100         105         110
Ser Gly Ser Phe Cys Arg Leu Ala Gly His Asp Ile Asn Tyr Leu Ala
115         120         125
Leu Ser Gly Val Leu Ser Lys Ile Gly Arg Ser Gly Glu Asn Pro Tyr
130         135         140
Ala Pro Leu Asn Leu Leu Ala Asp Phe Ala Gly Gly Gly Leu Met Cys

```

<210>	109
<211>	1524
<212>	DNA
<213>	Homo sapien
<400>	109
ggcacgagggc tgcgccaggg cctgagcgga ggcggggggca gcctcgccag cgggggcccc	60
gggcctggcc atgcctcact gagccagcgc ctgcgcctct acctcgcca cagctggaac	120
cagtgcgacc tagtggctct cacctgtctt ctctggggcg tgggetgcg gctgacctcg	180
ggtttgtacc acctgggccg cactgtcttc tgcatcgact tcattggttt caagggtgog	240
ctgcttcaca ttcttacggt caaaaaacag ctggggccca agatcgctcat cgtgagcaag	300
atgatgaagg acgtgttctt ctctctcttc ttctctggcg tgttgctggt agcctatggc	360
gtggccacgg aggggtctct gaggccacgg gacagtgacl tcccagaatl cctgcgcgc	420
gtctttctacc gtccctacct gcagatcttc gggcagattc ccacaggagga catggaagtg	480
gcctcatgg agcacagcaa ctgctcgctg gagcccggt tctgggcaca cctcctggg	540
gccaggcgg gcacctgogt ctcccaglat gccaaetggc tggtagtgct getcctcgte	600
atcttctctg tctggccaa catctgctg gtcaacttgc tcattgccat gttcagttac	660
acattcgga aagtaeagg caaeagegal ctctactgga aggcgacgg ttaccgccte	720
atccgggaat tccactcteg gcccgcgtg gcccgcctt ttatcgteat etcccaettg	780
cgctcctgct tcaggcaatt gtgcaggcga ccccgagcc cccagccgte etccccggc	840
ctcgagcatt tccgggttta cttttctaag gaagccgagc ggaagctgct aacgtgggaa	900

tcggtgcata	aggagaactt	tctgctggca	cgcgctaggg	acaagcggga	gagcgactcc	960
gagcgtctga	agcgcacgtc	ccagaaggtg	gacttggcac	tgaaacagct	gggacacatc	1020
cgcgagtacg	aacagcgctt	gaaagtgtctg	gagcgggagg	tccagcagtg	tagccgcgtc	1080
ctgggggtggg	tggccgaggg	cctgagccgc	tctgccttgc	tgcccccagg	tgggcccga	1140
ccccctgacc	tgcttgggtc	caaagactga	gccctgctgg	cggacttcaa	ggagaagccc	1200
ccacagggga	ttttgtctct	agagtaaggc	tcctctgggc	ctcggcccc	gcacctggtg	1260
gccttgtctt	tgaggtgagc	cccatgtcca	tctggggccac	tgtcaggacc	acctttggga	1320
gtgtcatcct	tacaaaccac	agcatgcccc	gctcctccca	gaaccagtcc	cagcctggga	1380
ggatcaaggc	ctggatcccc	ggccgttata	catctggagg	ctgcagggtc	cttggggtaa	1440
cagggaccac	agacccctca	ccactcacag	attcctcaca	ctggggaaat	aaagccattt	1500
cagaggaaaa	aaaaaaaaaa	aaaa				1524

<210> 110

<211> 3410

<212> DNA

<213> Homo sapien

<400> 110

gggaaccagc	ctgcacgcgc	tggctccggg	tgacagccgc	gcgcctcggc	caggatctga	60
gtgatgagac	gtgtccccac	tgaggtgccc	cacagcagca	ggtgttgagc	atgggctgag	120
aagctggacc	ggcaccaaag	ggctggcaga	aatgggcgcc	tggctgattc	ctaggcagtt	180
ggcggcagca	aggaggagag	gccgcagctt	ctggagcaga	gccgagacga	agcagttctg	240
gagtgcctga	acggcccccct	gagccctacc	cgccctggccc	actatggtcc	agaggctgtg	300
ggtgagccgc	ctgctgcggc	accggaaaagc	ccagctcttg	ctgggtcaacc	tgctaacctt	360
tggcctggag	gtgtgttttg	ccgcaggcat	cacctatgtg	ccgcctctgc	tgtctggaagt	420
gggggtagag	gagaagttca	tgaccatggt	gctgggcatt	ggtccagtgc	tgggcctggt	480
cttgttcccc	ctcctaggct	cagccagtga	ccactggcgt	ggacgctatg	gccgccgccg	540
gcccttcac	tgggcactgt	ccttgggcct	cctgctgagc	ctctttctca	tcccaagggc	600
cggctggcta	gcagggtctg	tgtgcccggg	tcccaggccc	ctggagctgg	cactgctcat	660
cctgggcgtg	gggctgctgg	acttctgttg	ccagggtgtg	ttcactccac	tggaggccct	720
gctctctgac	ctcttccggg	accgggacca	ctgtcgccag	gcctactctg	tctatgcctt	780
catgatcagt	cttgggggct	gcctgggcta	cctcctgcct	gccattgact	gggacaccag	840
tgccctggcc	ccctacctgg	gcacccaggga	ggagtgcctc	tttggcctgc	tcaccctcat	900
cttcttcacc	tgcgtagcag	ccacactgct	ggtggctgag	gaggcagcgc	tgggccccac	960
cgagccagca	gaagggtctg	cggccccctc	cttgtcgccc	cactgctgtc	catgccgggc	1020
ccgcttggct	ttccggaacc	tgggcgcctt	gcttccccgg	ctgcaccagc	tgtgctgccg	1080
catgccccgc	accctgcgcc	ggctcttcgt	ggctgagctg	tgcagctgga	tggcactcat	1140
gaccttcacg	ctgtttttaca	cggatttcgt	gggcgagggg	ctgtaccagg	gcgtgccccag	1200
agctgagccg	ggcaccgagg	cccggagaca	ctatgatgaa	ggcgttcgga	tgggcagcct	1260
ggggctgttc	ctgcagtgcg	ccatctccct	ggtcttctct	ctgggtcatg	accggctggt	1320
gcagcgattc	ggcactcgag	cagtctatct	ggccagtgtg	gcagctttcc	ctgtggctgc	1380
cggtgccaca	tgctgttccc	acagtgtggc	cgtgggtgaca	gcttcagccg	ccctcacccg	1440
gttaccttc	tcagccctgc	agatcctgcc	ctacacactg	gcctccctct	accaccggga	1500
gaagcaggtg	ttcctgcccc	aataccgagg	ggacactgga	ggtgctagca	gtgaggacag	1560
cctgatgacc	agcttcctgc	caggccctaa	gcctggagct	cccttcccta	atggacacgt	1620
gggtgctgga	ggcagtggcc	tgtccccacc	tccaccgcgc	ctctgcgggg	cctctgcctg	1680
tgatgtctcc	gtacgtgtgg	tgggtgggtga	gcccaccgag	gccagggtgg	ttccggggccg	1740
gggcatctgc	ctggacctcg	ccatcctgga	tagtgccttc	ctgctgtccc	aggtggcccc	1800
atccctgttt	atgggctcca	ttgtccagct	cagccagctc	gtcactgcct	atatggtgtc	1860
tgccgcaggc	ctgggtcttg	tcgccattta	ctttgtctaca	caggtagtat	ttgacaagag	1920
cgacttggcc	aaatactcag	cgtagaaaaa	ttccagcaca	ttgggggtgga	gggcctgcct	1980

cactgggtcc	cagctccccg	ctcctgttag	cccatggggg	ctgccgggct	ggccgccagt	2040
ttctgttgct	gccaaagtaa	tgtggctctc	tgctgccacc	ctgtgctgct	gaggtgctga	2100
gctgcacagc	tgggggctgg	ggcgtccctc	tcctctctcc	ccagtctcta	gggctgcctg	2160
actggaggcc	ttccaagggg	gtttcagttc	ggacttatac	agggaggcca	gaagggtccc	2220
atgactgga	atgcggggac	tctgcagggt	gattaccag	gctcagggtt	aacagctagc	2280
ctcctagtgt	agacacacct	agagaagggt	ttttgggagc	tgaataaact	cagtcacctg	2340
gtttcccatc	tctaagcccc	ttaacctgca	gcttcgttta	atgtagctct	tgcattgggag	2400
tttctaggat	gaaacactcc	tccatgggat	ttgaacatat	gacttatttg	taggggaaga	2460
gtcctgaggg	gcaacacaca	agaaccaggt	cccctcagcc	cacagcactg	tctttttgct	2520
gatccacccc	cctcttacct	tttatcagga	tgtggcctgt	tggtccttct	gttgccatca	2580
cagagacaca	ggcatttaaa	tatttaactt	atattattta	caaagtagaa	gggaatccat	2640
tgctagcttt	tctgtgttgg	tgtctaatat	ttgggtaggg	tgggggatcc	ccaacaatca	2700
ggccccctga	gatagctggg	cattgggctg	atcattgcca	gaatcttctt	ctcctggggg	2760
ctggcccccc	aaaatgccta	accagggacc	ttggaaattc	tactcatccc	aaatgataat	2820
tccaaatgct	gttacccaag	gttaggggtg	tgaaggaagg	tagaggggtg	ggcttcaggt	2880
ctcaacgggt	tccctaacca	cccctcttct	cttggcccag	cctgggtccc	cccacttcca	2940
ctccccctta	ctctctctag	gactgggctg	atgaaggcac	tgcccaaaat	ttccccctacc	3000
cccaactttc	ccctaccccc	aactttcccc	accagctcca	caaccctgtt	tggagctact	3060
gcaggaccag	aagcacaaag	tgcgggtttc	caagcctttg	tccatctcag	ccccagagt	3120
atatctgtgc	ttggggaatc	tcacacagaa	actcaggagc	accccctgcc	tgagctaagg	3180
gaggtcttat	ctctcagggg	gggtttaagt	gccgtttgca	ataatgtcgt	cttattttatt	3240
tagcggggtg	aataattttat	actgtaagtg	agcaatcaga	gtataatgtt	tatgggtgaca	3300
aaattaaagg	ctttctttata	tgttttaaaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	3360
aaaaaaaaara	aaaaaaaaaa	aaaaaaaaaa	aaaaaaataa	aaaaaaaaaa		3410

<210> 111

<211> 1289

<212> DNA

<213> Homo sapien

<400> 111

agccaggcgt	ccctctgcct	gccactcag	tggcaacacc	cgaggagctgt	tttgtccttt	60
gtggagcctc	agcagttccc	tctttcagaa	ctcactgcca	agagccctga	acaggagcca	120
ccatgcagtg	cttcagcttc	attaagacca	tgatgatcct	cttcaatttg	ctcatctttc	180
tgtgtggtgc	agccctgttg	gcagtgggca	tctgggtgtc	aatcgatggg	gcatcctttc	240
tgaagatctt	cgggccactg	tcgtccagtg	ccatgcagtt	tgtcaacgtg	ggctacttcc	300
tcacgcagc	cggcggttg	gtctttgtc	ttgggtttcct	gggctgctat	ggtgctaaga	360
ctgagagcaa	gtgtgccctc	gtgacgttct	tcttcaccc	cctcctcatc	ttcattgctg	420
aggttgcagc	tgctgtggtc	gccttggtgt	acaccacaat	ggctgagcac	ttcctgacgt	480
tgctggtagt	gcctgccatc	aagaaagatt	atgggttcca	ggaagacttc	actcaagtgt	540
ggaacaccac	catgaaagg	ctcaagtgt	gtggcttcac	caactatacg	gattttgagg	600
actcacccta	cttcaaagag	aacagtgcct	ttccccatt	ctgttgcaat	gacaacgtca	660
ccaacacagc	caatgaaacc	tgcaccaagc	aaaaggctca	cgacaaaaaa	gtagagggtt	720
gcttcaatca	gcttttgtat	gacatccgaa	ctaattgcag	caccgtgggt	ggtgtggcag	780
ctggaattgg	gggcctcgag	ctggctgcca	tgattgtgtc	catgtatctg	tactgcaatc	840
tacaataagt	ccacttctgc	ctctgccact	actgctgcca	catgggaact	gtgaagaggc	900
accctggcaa	gcagcagtg	ttgggggagg	ggacaggatc	taacaatgtc	acttgggcca	960
gaatggacct	gccctttctg	ctccagactt	ggggctagat	agggaccact	ccttttagcg	1020
atgcctgact	ttccttccat	tggtgggtgg	atgggtgggg	ggcattccag	agcctctaag	1080
gtagccagtt	ctgttgccca	ttccccaggt	ctattaaacc	cttgatatgc	cccctaggcc	1140
tagtggtgat	cccagtgctc	tactggggga	tgagagaaa	gcattttata	gcctgggcat	1200

<211> 553
 <212> PRT
 <213> Homo sapien

<400> 113

Met	Val	Gln	Arg	Leu	Trp	Val	Ser	Arg	Leu	Leu	Arg	His	Arg	Lys	Ala
1				5					10					15	
Gln	Leu	Leu	Leu	Val	Asn	Leu	Leu	Thr	Phe	Gly	Leu	Glu	Val	Cys	Leu
			20					25					30		
Ala	Ala	Gly	Ile	Thr	Tyr	Val	Pro	Pro	Leu	Leu	Leu	Glu	Val	Gly	Val
		35					40					45			
Glu	Glu	Lys	Phe	Met	Thr	Met	Val	Leu	Gly	Ile	Gly	Pro	Val	Leu	Gly
	50					55					60				
Leu	Val	Cys	Val	Pro	Leu	Leu	Gly	Ser	Ala	Ser	Asp	His	Trp	Arg	Gly
65					70				75						80
Arg	Tyr	Gly	Arg	Arg	Arg	Pro	Phe	Ile	Trp	Ala	Leu	Ser	Leu	Gly	Ile
			85						90					95	
Leu	Leu	Ser	Leu	Phe	Leu	Ile	Pro	Arg	Ala	Gly	Trp	Leu	Ala	Gly	Leu
			100					105					110		
Leu	Cys	Pro	Asp	Pro	Arg	Pro	Leu	Glu	Leu	Ala	Leu	Leu	Ile	Leu	Gly
		115					120					125			
Val	Gly	Leu	Leu	Asp	Phe	Cys	Gly	Gln	Val	Cys	Phe	Thr	Pro	Leu	Glu
	130					135					140				
Ala	Leu	Leu	Ser	Asp	Leu	Phe	Arg	Asp	Pro	Asp	His	Cys	Arg	Gln	Ala
145					150					155					160
Tyr	Ser	Val	Tyr	Ala	Phe	Met	Ile	Ser	Leu	Gly	Gly	Cys	Leu	Gly	Tyr
				165					170					175	
Leu	Leu	Pro	Ala	Ile	Asp	Trp	Asp	Thr	Ser	Ala	Leu	Ala	Pro	Tyr	Leu
			180					185					190		
Gly	Thr	Gln	Glu	Glu	Cys	Leu	Phe	Gly	Leu	Leu	Thr	Leu	Ile	Phe	Leu
	195						200					205			
Thr	Cys	Val	Ala	Ala	Thr	Leu	Leu	Val	Ala	Glu	Glu	Ala	Ala	Leu	Gly
	210					215					220				
Pro	Thr	Glu	Pro	Ala	Glu	Gly	Leu	Ser	Ala	Pro	Ser	Leu	Ser	Pro	His
225					230					235					240
Cys	Cys	Pro	Cys	Arg	Ala	Arg	Leu	Ala	Phe	Arg	Asn	Leu	Gly	Ala	Leu
				245					250					255	
Leu	Pro	Arg	Leu	His	Gln	Leu	Cys	Cys	Arg	Met	Pro	Arg	Thr	Leu	Arg
			260					265					270		
Arg	Leu	Phe	Val	Ala	Glu	Leu	Cys	Ser	Trp	Met	Ala	Leu	Met	Thr	Phe
		275					280					285			
Thr	Leu	Phe	Tyr	Thr	Asp	Phe	Val	Gly	Glu	Gly	Leu	Tyr	Gln	Gly	Val
	290				295						300				
Pro	Arg	Ala	Glu	Pro	Gly	Thr	Glu	Ala	Arg	Arg	His	Tyr	Asp	Glu	Gly
305					310					315					320
Val	Arg	Met	Gly	Ser	Leu	Gly	Leu	Phe	Leu	Gln	Cys	Ala	Ile	Ser	Leu
				325					330					335	
Val	Phe	Ser	Leu	Val	Met	Asp	Arg	Leu	Val	Gln	Arg	Phe	Gly	Thr	Arg
			340					345					350		
Ala	Val	Tyr	Leu	Ala	Ser	Val	Ala	Ala	Phe	Pro	Val	Ala	Ala	Gly	Ala
		355					360					365			

002230-002230

Ser Pro Tyr Phe Lys Glu Asn Ser Ala Phe Pro Pro Phe Cys Cys Asn
 165 170 175
 Asp Asn Val Thr Asn Thr Ala Asn Glu Thr Cys Thr Lys Gln Lys Ala
 180 185 190
 His Asp Gln Lys Val Glu Gly Cys Phe Asn Gln Leu Leu Tyr Asp Ile
 195 200 205
 Arg Thr Asn Ala Val Thr Val Gly Gly Val Ala Ala Gly Ile Gly Gly
 210 215 220
 Leu Glu Leu Ala Ala Met Ile Val Ser Met Tyr Leu Tyr Cys Asn Leu
 225 230 235 240
 Gln

<210> 115
 <211> 366
 <212> DNA
 <213> Homo sapien

<400> 115
 gctctttctc tcccctctc tgaatttaaat tctttcaact tgcaatttgc aaggattaca 60
 catttcactg tgatgtatat tgtgttgcaa aaaaaaaaaa gtgtctttgt ttaaaattac 120
 ttggtttgtg aatccatctt gctttttccc cattggaact agtcattaac ccatctctga 180
 actggtagaa aaacatctga agagctagtc tatcagcatc tgacagggtga attggatggt 240
 tctcagaacc atttcaccca gacagcctgt ttctatcctg ttttaataaat tagtttgggt 300
 tctctacatg cataacaaac cctgctccaa tctgtcacat aaaagtctgt gacttgaagt 360
 ttagtc 366

<210> 116
 <211> 282
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(282)
 <223> n = A,T,C or G

<400> 116
 acaaagatga accattttcct atattatagc aaaattaaaa tctaccgcgt ttctaattatt 60
 gagaaatgag atnaaacaca atnttataaa gtctacttag agaagatcaa gtgacctcaa 120
 agactttact attttcatat tttaagacac atgatttatc ctatttttagt aacctgggtc 180
 atacgttaaa caaaggataa tgtgaacagc agagaggatt tgttggcaga aaatctatgt 240
 tcaatctnga actatctana tcacagacat ttctattcct tt 282

<210> 117
 <211> 305
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature

<400> 120
 actcgttgca natcaggggc cccccagagt caccgttgca ggagtccttc tgggtcttgcc 60
 ctccgccggc gcagaacatg ctgggggtggt 90

<210> 121
 <211> 218
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(218)
 <223> n = A,T,C or G

<400> 121
 tgtancgtga anacgacaga naggggtgtc aaaaatggag aanccttgaa gtcattttga 60
 gaataagatt tgctaaaaga tttgggggcta aaacatgggt attgggagac atttctgaag 120
 atatncangt aaattangga atgaattcat ggttcctttg ggaattcctt tacgatngcc 180
 agcatanact tcatgtgggg atancagcta cccttgta 218

<210> 122
 <211> 171
 <212> DNA
 <213> Homo sapien

<400> 122
 taggggtgta tgcaactgta aggacaaaaa ttgagactca actggcttaa ccaataaagg 60
 catttgtag ctcatggaac aggaagtcgg atgggtggggc atcttcagtg ctgcatgagt 120
 caccaccccg gcgggggtcat ctgtgccaca ggtccctggt gacagtgcgg t 171

<210> 123
 <211> 76
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(76)
 <223> n = A,T,C or G

<400> 123
 tgtagcgtga agacnacaga atgggtgtgtg ctgtgctatc caggaacaca tttattatca 60
 ttatcaanta ttgtgt 76

<210> 124
 <211> 131
 <212> DNA
 <213> Homo sapien

<400> 124
 acctttcccc aaggccaatg tctgtgtgtg taactggccg gctgcaggac agctgcaatt 60

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(192)

<223> n = A,T,C or G

<400> 129

acatacatgt	gtgtatat	tttaa	atatca	ctttt	gtatc	actct	gactt	tttag	catac	60
tgaaaacaca	ctaacataat	ttntgtgaac	catgatcaga	tacaacccaa	atcattcatc					120
tagcacattc	atctgtgata	naaagatagg	tgagtttcat	ttccttcacg	ttggccaatg					180
gataaaca	aaa gt									192

<210> 130

<211> 362

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(362)

<223> n = A,T,C or G

<400> 130

ccctttttta	tggaatgagt	agactgtatg	tttgaanatt	tanccacaac	ctctttgaca	60
tataatgacg	caacaaaaag	gtgctgttta	gtcctatggg	tcagtttatg	cccctgacaa	120
gtttccattg	tgttttgccg	atcttctggc	taatcgtggg	atcctccatg	ttattagtaa	180
ttctgtattc	cattttgtta	acgcctggta	gatgtaacct	gctangaggc	taactttata	240
cttatttaaa	agctcttatt	ttgtgggtcat	taaaatggca	atttatgtgc	agcactttat	300
tgcagcagga	agcacgtgtg	ggttgggttg	aaagctcttt	gctaactcta	aaaagtaatg	360
gg						362

<210> 131

<211> 332

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(332)

<223> n = A,T,C or G

<400> 131

ctttttgaaa	gatcgtgtcc	actcctgtgg	acatcttggt	ttaatggagt	ttcccatgca	60
gtangactgg	tatggttgca	gctgtccaga	taaaaacatt	tgaagagctc	caaaatgaga	120
gttctcccag	gttcgccctg	ctgctccaag	tctcagcagc	agcctctttt	aggaggcatc	180
ttctgaacta	gattaaggca	gcttgtaaat	ctgatgtgat	ttgggtttatt	atccaactaa	240
cttccatctg	ttatcactgg	agaaagccca	gactcccan	gacnggtacg	gattgtgggc	300
atanaaggat	tgggtgaagc	tggcgttgtg	gt			332

<210> 132

<211> 322
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(322)
 <223> n = A,T,C or G

<400> 132
 acttttgcca ttttgtatat ataaacaatc ttgggacatt ctcttgaaaa ctaggtgtcc 60
 agtggcctaag agaactcgat ttcaagcaat tctgaaagga aaaccagcat gacacagaat 120
 ctcaaattcc caaacagggg ctctgtggga aaaatgaggg aggaccttg tatctcgggt 180
 tttagcaagt taaaatgaan atgacaggaa aggcttattt atcaacaaag agaagagttg 240
 ggatgcttct aaaaaaaact ttggtagaga aaataggaat gctnaatcct agggaagcct 300
 gtaacaatct acaattgggc ca 322

<210> 133
 <211> 278
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(278)
 <223> n = A,T,C or G

<400> 133
 acaagccttc acaagtttaa ctaaattggg attaatcttt ctgtanttat ctgcataatt 60
 cttgtttttc tttccatctg gctcctgggt tgacaatttg tggaaacaac tctattgcta 120
 ctatttaaaa aaaatcacaa atctttccct ttaagctatg ttnaattcaa actattcctg 180
 ctattcctgt tttgtcaaag aaattatatt tttcaaaata tgtntatttg tttgatgggt 240
 cccacgaaac actaataaaa accacagaga ccagcctg 278

<210> 134
 <211> 121
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(121)
 <223> n = A,T,C or G

<400> 134
 gtttanaaaa cttgttttagc tccatagagg aaagaatggt aaactttgta ttttaaaaca 60
 tgattctctg aggttaaact tgggttttcaa atgttatttt tacttgtatt ttgcttttgg 120
 t 121

<210> 135
 <211> 350

<212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(350)
 <223> n = A,T,C or G

<400> 135
 acttanaacc atgcctagca catcagaatc cctcaaagaa catcagtata atcctataacc 60
 atancaagtg gtgactgggt aagcgtgcga caaaggtcag ctggcacatt acttgtgtgc 120
 aaacttgata cttttgttct aagtaggaac tagtatacag tncctaggan tggtagtcca 180
 ggggtgcccc caactcctgc agccgctcct ctgtgccagn ccctgnaagg aactttcgct 240
 ccacctcaat caagccctgg gccatgctac ctgcaattgg ctgaacaaac gtttgctgag 300
 ttcccaagga tgcaaagcct ggtgctcaac tcctggggcg tcaactcagt 350

<210> 136
 <211> 399
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(399)
 <223> n = A,T,C or G

<400> 136
 tgtaccgtga agacgacaga agttgcatgg cagggacagg gcagggccga ggccagggtt 60
 gctgtgattg tatccgaata ntccctcgta gaaaagataa tgagatgacg tgagcagcct 120
 gcagacttgt gtctgccttc aanaagccag acaggaaggc cctgcctgcc ttggctctga 180
 cctggcggcc agccagccag ccacaggtgg gcttcttcct tttgtggtga caacnccaag 240
 aaaactgcag agggccaggg tcaggtgtna gtgggtangt gaccataaaa caccaggtgc 300
 tcccaggaac ccgggcaaag gccatcccca cctacagcca gcatgcccac tggcgtgatg 360
 ggtgcagang gatgaagcag ccagntgttc tgctgtggt 399

<210> 137
 <211> 165
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(165)
 <223> n = A,T,C or G

<400> 137
 actggtgtgg tnggggggtga tgctgggtgg anaagttgan gtgacttcan gatggtgtgt 60
 ggaggaagtg tgtgaacgta gggatgtaga ngttttggcc gtgctaaatg agcttcggga 120
 ttggctggtc ccactggtgg tcaactgtcat tgggtggggt cctgt 165

<210> 138

<220>
 <221> misc_feature
 <222> (1)...(335)
 <223> n = A,T,C or G

<400> 141
 actttatttt caaaacactc atatgttgca aaaaacacat agaaaaataa agtttggtgg 60
 gggtgctgac taaacttcaa gtcacagact tttatgtgac agattggagc aggggttggt 120
 atgcatgtag agaaccctaa ctaatttatt aaacaggata gaaacaggct gtctgggtga 180
 aatggttctg agaaccatcc aattcacctg tcagatgctg atanactagc tcttcagatg 240
 tttttctacc agttcagaga tnggttaatg actanttcca atggggaaaa agcaagatgg 300
 attcacaac caagtaattt taaacaaaga cactt 335

<210> 142
 <211> 459
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(459)
 <223> n = A,T,C or G

<400> 142
 accagggttaa tattgccaca tatatccttt ccaattgcgg gctaaacaga cgtgtattta 60
 gggttgttta aagacaacct agcttaatat caagagaaat tgtgacctt catggagtat 120
 ctgatggaga aaacactgag ttttgacaaa tcttatttta ttcagatagc agtctgatca 180
 cacatggtcc aacaacactc aaataataaa tcaaataatna tcagatgtta aagattggtc 240
 ttcaaacatc atagccaatg atgccccgct tgcctataat ctctccgaca taaaaccaca 300
 tcaacacctc agtggccacc aaaccattca gcacagcttc cttaactgtg agctgtttga 360
 agctaccagt ctgagcacta ttgactatnt ttttcangct ctgaatagct ctagggatct 420
 cagcangggg gggaggaacc agctcaacct tggcgtant 459

<210> 143
 <211> 140
 <212> DNA
 <213> Homo sapien

<400> 143
 acatttcctt ccaccaagtc aggactcctg gcttctgtgg gagttcttat cacctgaggg 60
 aaatccaaac agtctctcct agaaaggaat agtgtcacca accccaccca tctccctgag 120
 accatccgac ttccctgtgt 140

<210> 144
 <211> 164
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature

<222> (1)...(173)

<223> n = A,T,C or G

<400> 147

acattgtttt	tttgagataa	agcattgana	gagctctcct	taacgtgaca	caatggaagg	60
actggaacac	ataccacat	ctttgttctg	agggataatt	ttctgataaa	gtcttgctgt	120
atattcaagc	acatatgtta	tatattattc	agttccatgt	ttatagccta	ggt	173

<210> 148

<211> 477

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(477)

<223> n = A,T,C or G

<400> 148

acaaccactt	tatctcatcg	aatttttaac	ccaaactcac	tcactgtgcc	tttctatcct	60
atgggatata	ttatttgatg	ctccatttca	tcacacatat	atgaataata	cactcatact	120
gccctactac	ctgctgcaat	aatcacattc	ccttcctgtc	ctgaccctga	agccattggg	180
gtggtcctag	tggccatcag	tccangcctg	caccttgagc	ccttgagctc	cattgctcac	240
nccanccac	ctcaccgacc	ccatcctctt	acacagctac	ctccttgctc	tctaacccca	300
tagattatnt	ccaaattcag	tcaattaagt	tactattaac	actctaccgg	acatgtccag	360
caccactggt	aagccttctc	cagccaacac	acacacacac	acacncacac	acacacatat	420
ccaggcacag	gctacctcat	cttcacaatc	acccctttaa	ttaccatgct	atggtgg	477

<210> 149

<211> 207

<212> DNA

<213> Homo sapien

<400> 149

acagttgtat	tataatatca	agaaataaac	ttgcaatgag	agcattttaag	agggagaagac	60
taacgtatth	tagagagcca	aggaaggtht	ctgtggggag	tgggatgtaa	ggtggggcct	120
gatgataaat	aagagtcagc	caggtaagtg	ggtggtgtgg	tatgggcaca	gtgaagaaca	180
tttcaggcag	agggacacag	agtgaaa				207

<210> 150

<211> 111

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(111)

<223> n = A,T,C or G

<400> 150

accttgattt	cattgctgct	ctgatggaaa	cccaactatc	taatttagct	aaaacatggg	60
------------	------------	------------	------------	------------	------------	----

cacttaaattg tggtcagtgt ttggacttgt taactantgg catctttggg t 111

<210> 151
 <211> 196
 <212> DNA
 <213> Homo sapien

<400> 151
 agcgcggcag gtcattattga acattccaga tacctatcat tactcgatgc tgttgataac 60
 agcaagatgg ctttgaactc agggtcacca ccagctattg gaccttacta tgaaaaccat 120
 ggataccaac cggaaaaccc ctatcccgcg cagcccactg tggccccac tgtctacgag 180
 gtgcatccgg ctccagt 196

<210> 152
 <211> 132
 <212> DNA
 <213> Homo sapien

<400> 152
 acagcacttt cacatgtaag aaggagagaaa ttcctaaatg taggagaaag ataacagAAC 60
 cttccccctt tcatctagtg gtggaaacct gatgctttat gttgacagga atagaaccag 120
 gagggagtgt gt 132

<210> 153
 <211> 285
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(285)
 <223> n = A,T,C or G

<400> 153
 acaanaccca nganaggcca ctggccgtgg tgtcatggcc tccaaacatg aaagtgtcag 60
 cttctgctct tatgtcctca tctgacaact ctttaccatt tttatcctcg ctcagcagga 120
 gcacatcaat aaagtccaaa gtcttggact tggccttggc ttggaggaag tcatcaacac 180
 cctggctagt gaggggtgcg cgccgctcct ggatgacggc atctgtgaag tcgtgcacca 240
 gtctgcaggc cctgtggaag cgccgtccac acggagtnag gaatt 285

<210> 154
 <211> 333
 <212> DNA
 <213> Homo sapien

<400> 154
 accacagtcc tgttgggcca gggcttcatg accctttctg tgaaaagcca tattatcacc 60
 accccaaatt tttccttaaa tatctttaac tgaaggggtc agcctcttga ctgcaaagac 120
 cctaagccgg ttacacagct aactcccact ggccctgatt tgtgaaattg ctgctgcctg 180
 attggcacag gagtcaagg tgttcagctc cctcctccg tggaacgaga ctctgatttg 240
 agtttcacaa attctcgggc cacctcgtca ttgctcctct gaaataaaat ccggagaatg 300

gtcaggcctg tctcatccat atggatcttc cgg

333

<210> 155
 <211> 308
 <212> DNA
 <213> Homo sapien

 <220>
 <221> misc_feature
 <222> (1)...(308)
 <223> n = A,T,C or G

<400> 155
 actggaaata ataaaaccca catcacagtg ttgtgtcaaa gatcatcagg gcatggatgg 60
 gaaagtgctt tgggaactgt aaagtgccta acacatgata gatgattttt gttataatat 120
 ttgaatcacg gtgcatacaa actctcctgc ctgctcctcc tgggccccag cccagcccc 180
 atcacagctc actgctctgt tcatccaggc ccagcatgta gtggctgatt cttcttggct 240
 gcttttagcc tccanaagtt tctctgaagc caaccaaacc tctangtgta aggcattgctg 300
 gccctggt 308

<210> 156
 <211> 295
 <212> DNA
 <213> Homo sapien

<400> 156
 accttgctcg gtgcttggaa catattagga actcaaaata tgagatgata acagtgccta 60
 ttattgatta ctgagagAAC tgtagacat ttagttgaag attttctaca caggaaactga 120
 gaataggaga ttatgttttg cctcatatt ctctcctatc ctcttgcct cattctatgt 180
 ctaatatatt ctcaatcaaa taagggttagc ataatcagga aatcgaccaa ataccaatat 240
 aaaaccagat gtctatcctt aagattttca aatagaaaac aaattaacag actat 295

<210> 157
 <211> 126
 <212> DNA
 <213> Homo sapien

<400> 157
 acaagtttaa atagtgtgtg cactgtgcat gtgctgaaat gtgaaatcca ccacatttct 60
 gaagagcaaa acaaattctg tcatgtaatc tctatcttgg gtcgtgggta tatctgtccc 120
 cttagt 126

<210> 158
 <211> 442
 <212> DNA
 <213> Homo sapien

 <220>
 <221> misc_feature
 <222> (1)...(442)
 <223> n = A,T,C or G

<400> 158
 acccactggt ctiggaaca cccatcctta atacgatgat ttttctgtcg tgtgaaaatg 60
 aanccagcag gctgccccta gtcagtcctt ccttcagag aaaaagagat ttgagaaagt 120
 gcctgggtaa ttcaccatta atttcctccc ccaaactctc tgagtcttcc cttaatatTT 180
 ctggtggttc tgaccaaagc aggtcatggt ttgttgagca tttgggatcc cagtgaagta 240
 natgtttgta gccttgcata cttagccctt cccacgcaca aacggagtgg cagagtgggtg 300
 ccaaccctgt tttcccagtc cacgtagaca gattcacagt gcggaattct ggaagctgga 360
 nacagacggg ctctttgcag agccgggact ctgagangga catgagggcc tctgcctctg 420
 tgttcattct ctgatgtcct gt 442

<210> 159
 <211> 498
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(498)
 <223> n = A,T,C or G

<400> 159
 acttccaggt aacgttggtg tttccgttga gcctgaactg atgggtgacg ttgtaggttc 60
 tccaacaaga actgaggttg cagagcgggt agggaagagt gctgttccag ttgcacctgg 120
 gctgctgtgg actgttggtg attcctcact acggcccaag gttgtggaac tggcanaaaag 180
 gtgtgtgtgt gganttgagc tcgggcggct gtggtagggt gtgggctctt caacaggggc 240
 tgctgtggtg ccgggangtg aangtggtgt gtcacttgag cttggccagc tctggaaagt 300
 antanattct tcctgaaggc cagcgttgtt ggagctggca ngggtcantg ttgtgtgtaa 360
 cgaaccagtg ctgctgtggg tgggtgtana tcctccacaa agcctgaagt tatggtgtcn 420
 tcaghtaana atgtggtttc agtgtccctg ggcngctgtg gaaggttgta nattgtcacc 480
 aagggaataa gctgtggt 498

<210> 160
 <211> 380
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(380)
 <223> n = A,T,C or G

<400> 160
 acctgcatcc agcttcccctg ccaaactcac aaggagacat caacctctag acagggaaac 60
 agcttcagga tacttccagg agacagagcc accagcagca aaacaaatat tcccatgcct 120
 ggagcatggc atagaggaag ctganaaatg tggggctctga ggaagccatt tgagtctggc 180
 cactagacat ctcatcagcc acttgtgtga agagatgccc catgacccca gatgcctctc 240
 ccacccttac ctccatctca cacacttgag ctttccactc tgtataattc taacatcctg 300
 gagaaaaatg gcagtttgac cgaacctgtt cacaacggta gaggttgatt tctaacgaaa 360
 cttgtagaat gaagcctgga 380

tctagtaggc acagggctcc caggccaggc ctcatctctc tctggcctct aatagtcaat 420
gattgtgtag ccattgcctat cagtaaaaag atntttgagc aaacacttt 469

<210> 165
<211> 195
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(195)
<223> n = A,T,C or G

<400> 165
acagtttttt atanatatcg acattgccgg cacttggtgtt cagtttcata aagctgggtgg 60
atccgctgtc atccactatt ccttggttag agtaaaaatt attcttatag cccatgtccc 120
tgcaggccgc ccgcccgtag ttctcgttcc agtcgtcttg gcacacaggg tgccaggact 180
tcctctgaga tgagt 195

<210> 166
<211> 383
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(383)
<223> n = A,T,C or G

<400> 166
acatcttagt agtgtggcac atcagggggc catcagggtc acagtcactc atagcctcgc 60
cgaggtcgga gtccacacca ccggtgtagg tgtgctcaat cttgggcttg gcgcccacct 120
ttggagaagg gatatgctgc acacacatgt ccacaaagcc tgtgaactcg ccaaagaatt 180
tttgagacc agctgagca aggggcggat gttcagcttc agctcctcct tcgtcaggtg 240
gatgccaaac tcgtctangg tccgtgggaa gctgggtgtcc acntcaccta caacctgggc 300
gangatctta taaagaggct ccnagataaa ctccacgaaa cttctctggg agctgctagt 360
nggggccttt ttggtgaact ttc 383

<210> 167
<211> 247
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(247)
<223> n = A,T,C or G

<400> 167
acagagccag accttggcca taaatgaanc agagattaag actaaacccc aagtcganat 60
tgagacagaa actggagcaa gaagtgggcc tggggctgaa gtagagacca aggccactgc 120

004090:004090

tatanccata cacagagcca actctcaggc caaggcnatg gttggggcag anccagagac 180
tcaatctgan tccaaagtgg tggctggaac actgggtcatg acanaggcag tgactctgac 240
tganctc 247

<210> 168
<211> 273
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(273)
<223> n = A,T,C or G

<400> 168
acttctaagt tttctagaag tggaaggatt gtantcatcc tgaaaatggg tttacttcaa 60
aatccctcan ccttgttctt cacnactgtc tatactgana gtgtcatgtt tccacaaagg 120
gctgacacct gagcctgnat tttcactcat ccctgagaag ccctttccag taggggtggc 180
aattcccaac ttccttgcca caagcttccc aggctttctc ccctggaaaa ctccagcttg 240
agtcccagat acactcatgg gctgccttg gca 273

<210> 169
<211> 431
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(431)
<223> n = A,T,C or G

<400> 169
acagccttgg cttccccaaa ctccacagtc tcagtgcaga aagatcatct tccagcagtc 60
agctcagacc aggggtcaaag gatgtgacat caacagtttc tggtttcaga acagggttcta 120
ctactgtcaa atgaccccc atacttcctc aaaggctgtg gtaagttttg cacagggtgag 180
ggcagcagaa aggggggtant tactgatgga caccatcttc tctgtatact ccacactgac 240
cttgccatgg gcaaaggccc ctaccacaaa aacaatagga tcaactgctgg gcaccagctc 300
acgcacatca ctgacaaccg ggatggaaaa agaantgcc aactttcatac atccaactgg 360
aaagtgatct gatactggat tcttaattac cttcaaaagc ttctgggggc catcagctgc 420
tcgaacactg a 431

<210> 170
<211> 266
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(266)
<223> n = A,T,C or G

Met Val Glu Ala Ser Leu Ser Val Arg His Pro Glu Tyr Asn Arg Pro
 1 5 10 15
 Leu Leu Ala Asn Asp Leu Met Leu Ile Lys Leu Asp Glu Ser Val Ser
 20 25 30
 Glu Ser Asp Thr Ile Arg Ser Ile Ser Ile Ala Ser Gln Cys Pro Thr
 35 40 45
 Ala Gly Asn Ser Cys Leu Val Ser Gly Trp Gly Leu Leu Ala Asn Gly
 50 55 60
 Arg Met Pro Thr Val Leu Gln Cys Val Asn Val Ser Val Val Ser Glu
 65 70 75 80
 Glu Val Cys Ser Lys Leu Tyr Asp Pro Leu Tyr His Pro Ser Met Phe
 85 90 95
 Cys Ala Gly Gly Gly Gln Xaa Gln Xaa Asp Ser Cys Asn Gly Asp Ser
 100 105 110
 Gly Gly Pro Leu Ile Cys Asn Gly Tyr Leu Gln Gly Leu Val Ser Phe
 115 120 125
 Gly Lys Ala Pro Cys Gly Gln Val Gly Val Pro Gly Val Tyr Thr Asn
 130 135 140
 Leu Cys Lys Phe Thr Glu Trp Ile Glu Lys Thr Val Gln Ala Ser
 145 150 155

<210> 173

<211> 1265

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(1265)

<223> n = A,T,C or G

<400> 173

ggcagcccg	actgcagcc	ctggcaggcg	gcactgggtca	tggaaaacga	attgttctgc	60
tcggggtcc	tgggtcatcc	gcagtgggtg	ctgtcagccg	cacactgttt	ccagaactcc	120
tacaccatcg	ggctgggcct	gcacagtctt	gaggccgacc	aagagccagg	gagccagatg	180
gtggaggcca	gcctctccgt	acggcaccga	gagtacaaca	gacccttgct	cgctaacgac	240
ctcatgtctca	tcaagttgga	cgaatccgtg	tccgagtctg	acaccatccg	gagcatcagc	300
attgcttcgc	agtgccttac	cgcggggaac	tcttgccctg	tttctggctg	gggtctgctg	360
gcgaacggtg	agctcacggg	tgtgtgtctg	ccctcttcaa	ggaggtcctc	tgcccagtcg	420
cgggggctga	cccagagctc	tgcgtcccag	gcagaatgcc	taccgtgctg	cagtgcgtga	480
acgtgtcggg	ggtgtctgag	gaggtctgca	gtaagctcta	tgaccgctg	taccaccca	540
gcatgttctg	cgccggcgga	gggcaagacc	agaaggactc	ctgcaacggg	gactctgggg	600
ggccccctgat	ctgcaacggg	tacttgccag	gccttgtgtc	tttcggaaaa	gccccgtgtg	660
gccaagttgg	cgtgccaggt	gtctacacca	acctctgcaa	attcactgag	tggatagaga	720
aaaccgtcca	ggccagttaa	ctctggggac	tgggaaccca	tgaaattgac	ccccaaatac	780
atcctgcgga	aggaattcag	gaatatctgt	tcccagcccc	tcctccctca	ggcccaggag	840
tccaggcccc	cagcccctcc	tccctcaaac	caagggtaca	gatccccagc	ccctcctccc	900
tcagaccag	gagtccagac	ccccagcccc	ctcctccctc	agaccagga	gtccagcccc	960
tcctccntca	gaccaggag	tccagacccc	ccagccccc	ctccctcaga	cccaggggtt	1020
gaggccccca	accctcctc	cttcagagtc	agagggtocaa	gcccccaacc	cctcggtccc	1080
cagaccagga	ggtnnaggtc	ccagccccctc	tccntcaga	cccagnggtc	caatgccacc	1140


```

      35              40              45
Glu Ala Ser Leu Ser Val Arg His Pro Glu Tyr Asn Arg Pro Leu Leu
  50              55              60
Ala Asn Asp Leu Met Leu Ile Lys Leu Asp Glu Ser Val Ser Glu Ser
  65              70              75              80
Asp Thr Ile Arg Ser Ile Ser Ile Ala Ser Gln Cys Pro Thr Ala Gly
      85              90              95
Asn Ser Cys Leu Val Ser Gly Trp Gly Leu Leu Ala Asn Asp Ala Val
      100             105             110
Ile Ala Ile Gln Ser Xaa Thr Val Gly Gly Trp Glu Cys Glu Lys Leu
      115             120             125
Ser Gln Pro Trp Gln Gly Cys Thr Ile Ser Ala Thr Ser Ser Ala Arg
      130             135             140
Thr Ser Cys Cys Ile Leu Thr Gly Cys Ser Leu Leu Leu Thr Ala Ser
  145             150             155             160
Pro Gly Thr Leu

```

```

<210> 179
<211> 250
<212> DNA
<213> Homo sapien

```

```

<400> 179
ctggagtgcc ttggtgtttc aagcccctgc aggaagcaga atgcaccttc tgaggcacct      60
ccagctgccc ccggccgggg gatgcgaggc tcggagcacc cttgcccggc tgtgattgct      120
gccaggcact gttcatctca gcttttctgt ccctttgctc ccggcaagcg cttctgctga      180
aagttcatat ctggagcctg atgtcttaac gaataaaggt cccatgctcc acccgaaaaa      240
aaaaaaaaa                                     250

```

```

<210> 180
<211> 202
<212> DNA
<213> Homo sapien

```

```

<400> 180
actagtccag tgtggtggaa ttccattgtg ttgggcccac cacaatggct acctttaaca      60
tcacccagac cccgcccctg cccgtgcccc acgtgctgct taacgacagt atgatgctta      120
ctctgctact cggaaactat ttttatgtaa ttaatgtatg ctttcttggt tataaatgcc      180
tgatttaaaa aaaaaaaaaa aa                                     202

```

```

<210> 181
<211> 558
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(558)
<223> n = A,T,C or G

```


<220>
 <221> misc_feature
 <222> (1)...(496)
 <223> n = A,T,C or G

<400> 184
 accgaattgg gaccgctggc ttataagcga tcatgtyynt ccrgtatkac ctcaacgagc 60
 agggagatcg agtctatacg ctgaagaaat ttgacccgat gggacaacag acctgctcag 120
 cccatcctgc tcggttctcc ccagatgaca aatactctsg acaccgaatc accatcaaga 180
 aacgcttcaa ggtgctcatg acccagcaac cgcgccctgt cctctgaggg tcccttaaac 240
 tgatgtcttt tctgccacct gttacccctc ggagactccg taaccaaact ctteggactg 300
 tgagccctga tgcctttttg ccagccatac tctttggcat ccagtctctc gtggcgattg 360
 attatgcttg tgtgaggcaa tcatggtggc atcacccata aagggaacac atttgacttt 420
 tttttctcat attttaaatt actacmagaw tattwmagaw waaatgawtt gaaaaactst 480
 taaaaaaaaa aaaaaa 496

<210> 185
 <211> 384
 <212> DNA
 <213> Homo sapien

<400> 185
 gctggtagcc tatggcgkgg cccacggagg ggctcctgag gccacggrac agtgacttcc 60
 caagtatcyt ggcsgcgtc ttctaccgtc cctacctgca gatcttcggg cagattcccc 120
 aggaggacat ggacgtggcc ctcatggagc acagcaactg ytcgtcggag cccggcttct 180
 gggcacaccc tcctggggcc caggcgggca cctgcgtctc ccagtatgcc aactggcttg 240
 tgggtgctgct cctcgctcatc ttctgctcg tggccaacat cctgctggtc aacttgetca 300
 ttgccatgtt cagttacaca ttcgggcaaag tacagggcaa cagcgatctc tactgggaag 360
 gcgcagcgtt accgcctcat ccgg 384

<210> 186
 <211> 577
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(577)
 <223> n = A,T,C or G

<400> 186
 gagttagctc ctccacaacc ttgatgaggt cgtctgcagt ggcctctcgc ttcataccgc 60
 tnccatcgtc ataactgtagg tttgccacca cytcttgga tcttggggcg gcntaatatt 120
 ccaggaaact ctcaatcaag tcaccgtcga tgaaacctgt gggctggttc tgtcttcgc 180
 tcggtgtgaa aggatctccc agaaggagt ctcgatcttc cccacacttt tgatgacttt 240
 attgagtcga ttctgcatgt ccagcaggag gttgtaccag ctctctgaca gtgaggtcac 300
 cagccctatc atgccgttga mcgtgccgaa garcaccgag ccttgtgtgg gggkkgaagt 360
 ctacccaga ttctgatta ccagagagcc gtggcaaaag acattgacaa actcgcccag 420
 gtggaaaaaag amcamctcct ggargtgctn gccgctctc gtcmgttggt ggcagcgctw 480
 tccttttgac acacaaacaa gttaaaggca ttttcagccc ccagaaantt gtcacatcc 540
 aagatntcgc acagcactna tccagttggg attaaat 577

<210> 194
 <211> 392
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(392)
 <223> n = A,T,C or G

<400> 194
 gaacggctgg accttgccctc gcattgtgct tgctggcagg gaataccttg gcaagcagyt 60
 ccagtccgag cagccccaga ccgctgccgc ccgaagctaa gcctgcctct ggccttcccc 120
 tccgcctcaa tgcagaacca gtagtgggag cactgtgttt agagttaaga gtgaacactg 180
 tttgatttta cttgggaatt tcctctgtta tatagctttt cccaatgcta atttccaaac 240
 aacaacaaca aaataacatg tttgcctgtt aagttgtata aaagtaggtg attctgtatt 300
 taaagaaaat attactgtta catatactgc ttgcaatttc tgtatttatt gktnctstgg 360
 aaataaatat agttattaaa gggtgtcant cc 392

<210> 195
 <211> 502
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(502)
 <223> n = A,T,C or G

<400> 195
 ccsttkgagg ggtkaggkyc cagttyccga gtggaagaaa caggccagga gaagtgcgtg 60
 ccgagctgag gcagatgttc ccacagtgac ccccagagcc stgggstata gtytctgacc 120
 cctcncaagg aaagaccacs ttctggggac atgggctgga gggcaggacc tagaggcacc 180
 aaggggaaggc cccattccgg ggstgttccc cgaggaggaa gggaaggggc tctgtgtgcc 240
 ccccasgagg aagaggccct gagtcctggg atcagacacc ccttcacgtg tatccccaca 300
 caaatgcaag ctcaccaagg tcccctctca gtccccttcc stacaccctg amcggccact 360
 gscscacacc caccagagc acgccacccg ccatggggar tgtgctcaag gartcgcnng 420
 gcarcgtgga catctngtcc cagaaggggg cagaatctcc aatagangga ctgarcmstt 480
 gctnanaaaa aaaaaaaaaa aa 502

<210> 196
 <211> 665
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(665)
 <223> n = A,T,C or G

<400> 196

```

ggttacttgg tttcattgcc accacttagt ggatgtcatt tagaaccatt ttgtctgctc      60
cctctggaag ccttgccgag agcggacttt gtaattgttg gagaataact gctgaatttt      120
wagctgtttk gagttgatts gcaccactgc acccacaact tcaatatgaa aacyawttga      180
actwatttat tatcttgtga aaagtataac aatgaaaatt ttgttcatac tgtattkac      240
aagtatgatg aaaagcaawa gatatatatt cttttattat gttaaattat gattgccatt      300
attaatcggc aaaatgtgga gtgtatgttc ttttcacagt aatatatgcc ttttgtaact      360
tcacttgggt attttattgt aaatgartta caaaattctt aatttaagar aatggatgt      420
watatttatt tcattaattt ctttcctkgt ttacgtwaat tttgaaaaga wtgcatgatt      480
tcttgacaga aatcgatctt gatgctgtgg aagtagtttg acccacatcc ctatgagttt      540
ttcttagaat gtataaaggt tgtagcccat cnaacttcaa agaaaaaaat gaccacatac      600
tttgcaatca ggctgaaatg tggcatgctn ttctaattcc aactttataa actagcaaan      660
aagtg                                         665

```

```

<210> 197
<211> 492
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(492)
<223> n = A,T,C or G

```

```

<400> 197
tttntttttt ttttttttgc aggaaggatt ccatttattg tggatgcatt ttcacaatat      60
atgtttattg gagcgatcca ttatcagtga aaagtatcaa gtgtttataa natttttagg      120
aaggcagatt cacagaacat gctngtcngc ttgcagtttt acctcgtana gatnacagag      180
aattatagtc naaccagtaa acnaggaatt tacttttcaa aagattaaat ccaaactgaa      240
caaaattcta ccctgaaact tactccatcc aaatattgga ataanagtca gcagtgatac      300
attctcttct gaacttttaga ttttctagaa aaatatgtaa tagtgatcag gaagagctct      360
tgttcaaaag tacaacnaag caatgttccc ttaccatagg ccttaattca aactttgatc      420
catttcactc ccatcacggg agtcaatgct acctgggaca cttgtatttt gttcatnctg      480
ancntggctt aa                                         492

```

```

<210> 198
<211> 478
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(478)
<223> n = A,T,C or G

```

```

<400> 198
tttnttttgn atttcantct gtannaanta ttttcattat gtttattana aaaatatnaa      60
tgtntccaen acaaatcatn ttacntnagt aagaggccan ctacattgta caacatacac      120
tgagtatatt ttgaaaagga caagttttaa gtanacncat attgccganc atancacatt      180
tatacatggc ttgattgata tttagcacag canaaactga gtgagttacc agaaanaaat      240
natatatgtc aatcngatth aagatacaaa acagatccta tggtaacatan catcntgtag      300
gagttgtggc tttatgttta ctgaaagtca atgcagttcc tgtacaaaga gatggccgta      360

```

```

agcattctag tacctctact ccatgggttaa gaatcgtaca cttatgttta catatgtnca    420
gggtaagaat tgtgttaagt naanttatgg agagggtccan gagaaaaatt tgatncaa    478

```

```

<210> 199
<211> 482
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(482)
<223> n = A,T,C or G

```

```

<400> 199
agtgacttgt cctccaacaa aaccccttga tcaagtttgt ggcactgaca atcagaccta    60
tgctagttcc tgtcatctat tcgctactaa atgcagactg gaggggacca aaaaggggca    120
tcaactccag ctggattatt ttggagcctg caaatctatt cctacttgta cggactttga    180
agtgattcag tttcctctac ggatgagaga ctggctcaag aatatcctca tgcagcttta    240
tgaagccnac tctgaacacg ctggttatct nagatgagaa ncagagaaat aaagtcnaga    300
aaatttacct ggangaagag aggccttngg ctggggacca tccattgaa ccttctctta    360
anggacttta agaanaaaact accacatgtn tgtngtatcc tgggtgccngg ccgtttantg    420
aacntngacn ncacccttnt ggaatanant cttgaacngn tctgaactt gctcctctgc    480
ga

```

```

<210> 200
<211> 270
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(270)
<223> n = A,T,C or G

```

```

<400> 200
cggccgcaag tgcaactcca gctggggccg tgcggacgaa gattctgcca gcagttggtc    60
cgactgcgac gacggcgccg gcgacagtcg caggtgcagc gcgggcccct ggggtcttgc    120
aaggctgagc tgacgccgca gaggtcgtgt cacgtcccac gaccttgacg ccgtcgggga    180
cagccggaac agagcccggg gaangcggga ggcctcgggg agcccctcgg gaagggcggc    240
ccgagagata cgcaggtgca ggtggccgcc

```

```

<210> 201
<211> 419
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(419)
<223> n = A,T,C or G

```

<400> 201

tttttttttt	ttttggaatc	tactgcgagc	acagcaggtc	agcaacaagt	ttattttgca	60
gctagcaagg	taacagggta	gggcatgggt	acatgttcag	gtcaacttcc	tttgtcgtgg	120
ttgattgggt	tgtctttatg	ggggcggggg	ggggtagggg	aaancgaagc	anaantaaca	180
tggagtgggt	gcacctccc	tgtagaacct	ggttacnaaa	gcttggggca	gttcacctgg	240
tctgtgaccg	tcattttctt	gacatcaatg	ttattagaag	tcaggatata	ttttagagag	300
tccactgtnt	ctggagggag	attaggggtt	cttgccaana	tccaancaaa	atccacntga	360
aaaagtggga	tgatncangt	acngaatacc	ganggcatan	ttctcatant	cggtggcca	419

<210> 202

<211> 509

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(509)

<223> n = A,T,C or G

<400> 202

tttntttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	60
tggcacttaa	tccattttta	tttcaaaatg	tctacaaant	ttnaatncnc	cattatacng	120
gtnattttnc	aaaatctaaa	nnttattcaa	atntnagcca	aantccttac	ncaaatnnaa	180
tacncncaaa	aatcaaaaat	atacntntct	ttcagcaaac	ttngttacat	aaattaaana	240
aatatatacg	gctggtgttt	tcaaagtaca	attatcttaa	cactgcaaac	atnttttnaa	300
ggaactaaaa	taaaaaaaaa	cactnccgca	aagggttaaag	ggaacaacaa	attcntttta	360
caacancnnc	nattataaaa	atcatatctc	aaatcttagg	ggaatatata	cttcacacng	420
ggatcttaac	ttttactnca	ctttgtttat	ttttttanaa	ccattgtntt	gggccaaca	480
caatggnaat	nccnccncc	tggtactagt				509

<210> 203

<211> 583

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(583)

<223> n = A,T,C or G

<400> 203

tttttttttt	ttttttttga	ccccctctt	ataaaaaaca	agttaccatt	ttattttact	60
tacacatatt	tattttataa	ttggtattag	atattcaaaa	ggcagctttt	aaaatcaaac	120
taaatggaaa	ctgccttaga	tacataattc	ttaggaatta	gcttaaaatc	tgccataaagt	180
gaaaatcttc	tctagctctt	ttgactgtaa	attttttgact	cttgtaaaac	atccaaattc	240
atttttcttg	tctttaaaat	tatctaattc	ttccattttt	tccctattcc	aagtcaattt	300
gcttctctag	cctcatttcc	tagctcttat	ctactattag	taagtggctt	ttttcctaaa	360
agggaaaaaca	ggaagagana	atggcacaca	aaacaaacat	tttatattca	tattttctacc	420
tacgttaata	aaatagcatt	ttgtgaagcc	agctcaaaaag	aaggcttaga	tccttttatg	480
tccatttttag	tcactaaacg	atatcnaaag	tgccagaatg	caaaagggtt	gtgaacattt	540
attcaaaagc	taatataaga	tatttcacat	actcatcttt	ctg		583

<400> 204

<400> 205

<220>

<221> misc_feature
 <222> (1)...(487)
 <223> n = A,T,C or G

<400> 206

```

tttttttttt ttttttagtc aagtttctna tttttattat aattaaagtc ttggtcattt    60
catttattag ctctgcaact tacatattta aattaaagaa acgttnttag acaactgtna    120
caatttataa atgtaagggtg ccattattga gtanatatat tcctccaaga gtggatgtgt    180
cccttctccc accaactaat gaancagcaa cattagttta attttattag tagatnatac    240
actgctgcaa acgctaattc tcttctccat ccccatgtng atattgtgta tatgtgtgag    300
ttggtnagaa tgcatcanca atctnacaat caacagcaag atgaagctag gcntgggctt    360
tcggtgaaaa tagactgtgt ctgtctgaat caaatgatct gacctatcct cggtggcaag    420
aactcttcga accgcttcct caaaggcngc tgccacattt gtggcntctn ttgcacttgt    480
ttcaaaa                                           487

```

<210> 207
 <211> 332
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(332)
 <223> n = A,T,C or G

<400> 207

```

tgaattgggt aaaagactgc atttttanaa ctagcaactc ttatttcttt cctttaaaaa    60
tacatagcat taaatcccaa atcctattta aagacctgac agcttgagaa ggtcactact    120
gcatttatag gaccttctgg tggttctgct gttacntttg aantctgaca atccttgana    180
atctttgcat gcagaggagg taaaagggtat tggattttca cagaggaana acacagcgca    240
gaaatgaagg ggccaggctt actgagcttg tccactggag ggctcatggg tgggacatgg    300
aaaagaaggc agcctaggcc ctggggagcc ca                                           332

```

<210> 208
 <211> 524
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(524)
 <223> n = A,T,C or G

<400> 208

```

agggcggtgt gcgaggggcg ttactgtttt gtctcagtaa caataaatac aaaaagactg    60
gttgtgttcc ggccccatcc aaccacgaag ttgatttctc ttgtgtgcag agtgactgat    120
tttaaaggac atggagcttg tcacaatgtc acaatgtcac agtgtgaagg gcacactcac    180
tcccgcgtga ttcacattta gcaaccaaca atagctcatg agtcatact tgtaaatact    240
tttggcagaa tacttnttga aacttgacaga tgataactaa gatccaagat atttcccaa    300
gtaaatagaa gtgggtcata atattaatta cctgttcaca tcagcttcca tttacaagtc    360
atgagcccag acactgacat caaactaagc ccacttagac tcctcaccac cagtctgtcc    420

```

tgatcatcaga caggaggctg tcaccttgac caaattctca ccagtcaatc atctatccaa 480
aaaccattac ctgatccact tccggtaatg caccaccttg gtga 524

<210> 209
<211> 159
<212> DNA
<213> Homo sapien

<400> 209
gggtgaggaa atccagagtt gccatggaga aaattccagt gtcagcattc ttgctccttg 60
tggccctctc ctacactctg gccagagata ccacagtcaa acctggagcc aaaaaggaca 120
caaaggactc tcgacccaaa ctgccccaga ccctctcca 159

<210> 210
<211> 256
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(256)
<223> n = A,T,C or G

<400> 210
actccctggc agacaaaggc agaggagaga gctctgttag ttctgtgttg ttgaactgcc 60
actgaatttc tttccacttg gactattaca tgccanttga gggactaatg gaaaaacgta 120
tggggagatt ttanccaatt tangtntgta aatggggaga ctggggcagg cgggagagat 180
ttgcagggtg naaatgggan ggctgggttg ttanatgaac agggacatag gaggtaggca 240
ccaggatgct aaatca 256

<210> 211
<211> 264
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(264)
<223> n = A,T,C or G

<400> 211
acattgtttt tttgagataa agcattgaga gagctctcct taacgtgaca caatggaagg 60
actggaacac ataccacat ctttgttctg agggataatt ttctgataaa gtcttgctgt 120
atattcaagc acatatgtta tatattattc agttccatgt ttatagccta gttaaggaga 180
ggggagatac attcngaaag aggactgaaa gaaatactca agtnggaaaa cagaaaaaga 240
aaaaaaggag caaatgagaa gcct 264

<210> 212
<211> 328
<212> DNA
<213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(328)
 <223> n = A,T,C or G

<400> 212
 acccaaaaat ccaatgctga atatttggct tcattattcc canattcttt gattgtcaaa 60
 ggatttaatg ttgtctcagc ttgggcactt cagttaggac ctaaggatgc cagccggcag 120
 gtttatatat gcagcaacaa tattcaagcg cgacaacagg ttattgaact tgcccgccag 180
 ttnaatttca ttcccattga cttgggatcc ttatcatcag ccagagagat tgaaaattta 240
 cccctacnac tctttactct ctgganaggg ccagtgggtg tagctataag cttggccaca 300
 tttttttttc cttttattcct ttgtcaga 328

<210> 213
 <211> 250
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(250)
 <223> n = A,T,C or G

<400> 213
 acttatgagc agagcgacat atccnagtgt agactgaata aaactgaatt ctctccagtt 60
 taaagcattg ctactgaag ggatagaagt gactgccagg agggaaagta agccaaggct 120
 cattatgcca aagganatat acatttcaat tctccaaact tcttcctcat tccaagagtt 180
 ttcaatattt gcatgaacct gctgataanc catgttaana aacaaatata tctctnacct 240
 tctcatcggt 250

<210> 214
 <211> 444
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(444)
 <223> n = A,T,C or G

<400> 214
 acccagaatc caatgctgaa tatttggctt cattattccc agattctttg attgtcaaag 60
 gatttaaatgt tgtctcagct tgggcacttc agttaggacc taaggatgcc agccggcag 120
 tttatatatg cagcaacaat attcaagcgc gacaacaggt tattgaactt gcccgccagt 180
 tgaatttcat tcccattgac ttgggatcct tatcatcagc canagagatt gaaaatttac 240
 ccctacgact ctttactctc tggagagggc cagtgggtgg agctataagc ttggccacat 300
 ttttttttcc tttattcctt tgctcagagat gcgattcatc catatgctan aaaccaacag 360
 agtgactttt acaaaaattcc tataganatt gtgaataaaa ccttacctat agttgccatt 420
 actttgctct ccctaataata cctc 444

```
<220>
<221> misc_feature
<222> (1)...(366)
<223> n = A,T,C or G
```

```
<210> 216
<211> 260
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(260)
<223> n = A,T,C or G
```

```
<210> 217
<211> 262
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(262)
<223> n = A,T,C or G
```

<400> 217						
acctacgtgg	gtaagtttan	aaatgttata	atttcaggaa	naggaacgca	tataattgta	60
tcttgctat	aattttctat	tttaataagg	aaatagcaaa	ttgggggtggg	gggaatgtag	120
ggcattctac	agtttgagca	aaatgcaatt	aaatgtggaa	ggacagcact	gaaaaatttt	180
atgaataatc	tgtatgatta	tatgtctcta	gagtagattt	ataattaqcc	acttacccta	240

atatccttca tgcttgtaaa gt

262

<210> 218

<211> 205

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(205)

<223> n = A,T,C or G

<400> 218

accaaagtgg tgcattaccg gaantggatc aangacacca tcgtggccaa cccctgagca	60
cccctatcaa ctcccttttg tagtaaaactt ggaaccttgg aaatgaccag gccaagactc	120
aggctcccc agttctactg acctttgtcc ttangtntna ngtcagggt tgctaggaaa	180
anaaatcagc agacacaggt gtaaa	205

<210> 219

<211> 114

<212> DNA

<213> Homo sapien

<400> 219

tactgttttg tctcagtaac aataaatata aaaagactgg ttgtgttccg gccccatcca	60
accacgaagt tgatttctct tgtgtgcaga gtgactgatt ttaaaggaca tgga	114

<210> 220

<211> 93

<212> DNA

<213> Homo sapien

<400> 220

actagccagc acaaaaggca gggtagcctg aattgctttc tgctctttac atttctttta	60
aaataagcat ttagtgctca gtccctactg agt	93

<210> 221

<211> 167

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(167)

<223> n = A,T,C or G

<400> 221

actangtgca ggtgcgcaca aatatttgtc gatattccct tcatcttgga ttccatgagg	60
tcttttgccc agcctgtggc tctactgtag taagtttctg ctgatgagga gccagnatgc	120
ccccactac cttccctgac gtcgccana aatcacccaa cctctgt	167

<210> 222
 <211> 351
 <212> DNA
 <213> Homo sapien

<400> 222
 agggcgtggt gcggagggcg gtactgacct cattagtagg aggatgcatt ctggcacccc 60
 gttcttcacc tgtcccccac tccttaaaag gccatactgc ataaagtcaa caacagataa 120
 atgtttgctg aattaaagga tggatgaaaa aaattaataa tgaatttttg cataatccaa 180
 ttttctcttt tatatttcta gaagaagttt ctttgagcct attagatccc gggaatcttt 240
 taggtgagca tgattagaga gcttgtagggt tgcttttaca tatatctggc atatttgagt 300
 ctcgtatcaa aacaatagat tggtaaagggt ggtattattg tattgataag t 351

<210> 223
 <211> 383
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(383)
 <223> n = A,T,C or G

<400> 223
 aaaacaaaca aacaaaaaaaa acaattcttc attcagaaaa attatcttag ggactgatat 60
 tggtaattat ggtcaattta atwrtrttkt ggggcatttc cttacattgt cttgacaaga 120
 ttaaaatgtc tgtgccaaaa ttttgtattt tatttgagga cttcttatca aaagtaatgc 180
 tgccaaagga agtctaagga attagtagtg ttcccmtcac ttgtttggag tgtgctattc 240
 taaaagattt tgatttctctg gaatgacaat tatattttta ctttggtggg ggaaanagtt 300
 ataggaccac agtcttcact tctgatactt gttaaattaat cttttattgc acttgttttg 360
 accattaagc tatatgttta aaa 383

<210> 224
 <211> 320
 <212> DNA
 <213> Homo sapien

<400> 224
 cccctgaagg cttcttggtta gaaaatagta cagttacaac caataggaac aacaaaaaga 60
 aaaagtttgt gacattgtag tagggagtgt gtacccttca ctcccatca aaaaaaaaaat 120
 ggatacatgg ttaaaggata raagggaat attttatcat atgttctaaa agagaaggaa 180
 gagaaaatac tactttctcr aaatggaagc ccttaaagggt gctttgatac tgaaggacac 240
 aaatgtggcc gtccatcttc ctttaragtt gcatgacttg gacacggtaa ctgttgagcgt 300
 tttaractcm gcattgtgac 320

<210> 225
 <211> 1214
 <212> DNA
 <213> Homo sapien

<400> 225

```
<210> 226
<211> 119
<212> DNA
<213> Homo sapien
```

```
<210> 227
<211> 818
<212> DNA
<213> Homo sapien
```

<400> 227						
acaattcata	gggacgacca	atgaggacag	ggaatgaacc	cggctctccc	ccagccctga	60
tttttgctac	atatggggtc	ccttttcatt	ctttgcaaaa	acactgggtt	ttctgagaac	120
acggacggtt	cttagcacia	tttgtgaaat	ctgtgtaraa	ccgggctttg	caggggagat	180
aattttcctc	ctctggagga	aaggtggtga	ttgacaggca	gggagacagt	gacaaggcta	240
gagaaaagcca	cgctcggcct	tctctgaacc	aggatggaac	ggcagacccc	tgaaaacgaa	300
gcttgtcccc	ttccaatcag	ccacttctga	gaacccccat	ctaacttcct	actggaaaag	360
agggcctcct	caggagcagt	ccaagagttt	tcaaagataa	cgtgacaact	accatctaga	420
ggaaaggggtg	caccctcagc	agagaagccg	agagcttaac	tctggtcgtt	tccagagaca	480
acctgctggc	tgtcttgga	tgcgccagc	ctttgagagg	ccactacccc	atgaacttct	540
gccatccact	ggacatgaag	ctgaggacac	tgggcttcaa	cactgagttg	tcatgagagg	600
gacaggctct	gccctcaagc	cggctgaggg	cagcaaccac	tctcctcccc	tttctcacgc	660
aaagccattc	ccacaaatcc	agaccatacc	atgaagcaac	gagacccaaa	cagtttggct	720
caagaggata	tgaggactgt	ctcagcctgg	ctttgggctg	acaccatgca	cacacacaag	780
gtccacttct	aggttttcag	cctagatggg	agtctgt			818

<210> 228
 <211> 744
 <212> DNA
 <213> Homo sapien

<400> 228
 actggagaca ctgttgaact tgatcaagac ccagaccacc ccaggtctcc ttcgtgggat 60
 gtcattgacgt ttgacatacc tttggaacga gcctcctcct tggaagatgg aagaccgtgt 120
 tcgtggccga cctggcctct cctggcctgt ttcttaagat gcggagtcac atttcaatgg 180
 taggaaaagt ggcttcgtaa aatagaagag cagtcactgt ggaactacca aatggcgaga 240
 tgctcgtgtc acattggggg gctttgggat aaaagattta tgagccaact attctctggc 300
 accagattct aggccagttt gttccactga agcttttccc acagcagtcc acctctgcag 360
 gctggcagct gaatggcctt cgggtggctc tgtggcaaga tcacactgag atcgatgggt 420
 gagaaggcta ggatgcttgt ctagtgcttct tagctgtcac gttggctcct tccaggttgg 480
 ccagacggtg ttggccactc ccttctaaaa cacaggcgcc ctcttggtga cagtgacctg 540
 ccgtggtatg ccttggccca ttccagcagt cccagttatg catttcaagt ttgggggttg 600
 ttcttttcgt taatgttctt ctgtgttggt agctgtcttc atttcctggg ctaagcagca 660
 ttggggagatg tggaccagag atccactcct taagaaccag tggcgaaaga cactttcttt 720
 cttcactctg aagtagctgg tggt 744

<210> 229
 <211> 300
 <212> DNA
 <213> Homo sapien

<400> 229
 cgagtctggg ttttgtctat aaagtttgat ccctcctttt ctcatccaaa tcatgtgaac 60
 cattacacat cgaaataaaa gaaagggtgg agacttgccc aacgccaggc tgacatgtgc 120
 tgcagggttg ttgtttttta attattattg ttagaaacgt caccacagc ccctgttaat 180
 ttgtatgtga cagccaactc tgagaagggt ctatttttcc acctgcagag gatccagtct 240
 cactagggtc ctcttgccc tcacactgga gtctccgcca gtgtgggtgc ccactgacat 300

<210> 230
 <211> 301
 <212> DNA
 <213> Homo sapien

<400> 230
 cagcagaaca aatacaata tgaagagtgc aaagatctca taaaatctat gctgaggaat 60
 gagcgacagt tcaaggagga gaagcttgca gagcagctca agcaagctga ggagctcagg 120
 caatataaag tcctgggttca cactcaggaa cgagagctga cccagttaag ggagaagtgt 180
 cggaaggga gagatgcctc cctctcattg aatgagcatc tccaggccct cctcactccg 240
 gatgaaccgg acaagtccca ggggcaggac ctccaagaaa cagacctcgg ccgcgaccac 300
 g 301

<210> 231
 <211> 301
 <212> DNA
 <213> Homo sapien

<400> 231

gcaagcacgc	tggcaaactct	ctgtcagggtc	agctccagag	aagccattag	tcatttttagc	60
caggaactcc	aagtccacat	ccttggcaac	tggggacttg	cgcaggttag	ccttgaggat	120
ggcaacacgg	gactttctcat	caggaagtgg	gatgtagatg	agctgatcaa	gacggccagg	180
tctgaggatg	gcaggatcaa	tgatgtcagg	cgggttggtg	ccgccaatga	tgaacacatt	240
tttttttg	gacatgccat	ccattttctgt	caggatctgg	ttgatgactc	ggtcagcagc	300
c						301

<210> 232

<211> 301

<212> DNA

<213> Homo sapien

<400> 232

agtaggtatt	tcgtgagaag	ttcaacacca	aaactggaac	atagttctcc	ttcaagtgtt	60
ggcgacagcg	gggttctctg	attctggaat	ataactttgt	gtaaattaac	agccacctat	120
agaagagtcc	atctgtctgtg	aaggagagac	agagaactct	gggttccgtc	gtcctgtcca	180
cgtgctgtac	caagtgtctgg	tgccagcctg	ttacctgttc	tactgaaaa	tctggctaata	240
gctcttgtgt	atcacttctg	attctgacaa	tcaatcaatc	aatggcctag	agcactgact	300
g						301

<210> 233

<211> 301

<212> DNA

<213> Homo sapien

<400> 233

atgactgact	ttccagtaag	gctctctaag	gggtaagtag	gaggatccac	aggatttgag	60
atgctaaggc	cccagagatc	gtttgatcca	accctcttat	tttcagaggg	gaaaatgggg	120
cctagaagtt	acagagcatc	tagctggtgc	gctggcacc	ctggcctcac	acagactccc	180
gagtagctgg	gactacaggc	acacagtcac	tgaagcaggc	cctggttagca	attctatgcg	240
tacaaattaa	catgagatga	gtagagactt	tattgagaaa	gcaagagaaa	atcctatcaa	300
c						301

<210> 234

<211> 301

<212> DNA

<213> Homo sapien

<400> 234

aggtcctaca	catcgagact	catccatgat	tgatatgaat	ttaaaaatta	caagcaaaga	60
cattttattc	atcatgatgc	tttcttttgt	ttcttctttt	cgttttcttc	tttttctttt	120
tcaatttcag	caacatactt	ctcaatttct	tcaggattta	aaatcttgag	ggattgatct	180
cgcctcatga	cagcaagtgc	aatgtttttg	ccacctgact	gaaccacttc	caggagtgcc	240
ttgatcacca	gcttaatggg	cagatcatct	gcttcaatgg	cttcgtcagt	atagttcttc	300
t						301

<210> 235

<211> 283

<212> DNA

<213> Homo sapien

<400> 235

tggtgctgtg	catcaggcgg	gtttgagaaa	tattcaattc	tcagcagaag	ccagaatttg	60
aattccctca	tcttttaggg	aatcatttac	caggtttgga	gaggattcag	acagctcagg	120
tgctttcact	aatgtctctg	aacttctgtc	cctctttgtt	catggatagt	ccaataaata	180
atgttatctt	tgaactgatg	ctcataggag	agaatataag	aactctgagt	gatatcaaca	240
ttagggattc	aaagaaatat	tagatttaag	ctcacactgg	tca		283

<210> 236

<211> 301

<212> DNA

<213> Homo sapien

<400> 236

aggtcctcca	ccaactgcct	gaagcacggg	taaaattggg	aagaagtata	gtgcagcata	60
aatactttta	aatcgatcag	atttcocctaa	cccacatgca	atcttcttca	ccagaagagg	120
tcggagcagc	atcattaata	ccaagcagaa	tgctgaatag	ataaatacaa	tggtatatag	180
tggttagacg	gcttcatgag	tacagtgtac	tgtgggtatcg	taatctggac	ttgggttgta	240
aagcatcgtg	taccagtcag	aaagcatcaa	tactcgacat	gaacgaatat	aaagaacacc	300
a						301

<210> 237

<211> 301

<212> DNA

<213> Homo sapien

<400> 237

cagtggtagt	ggtgggtggac	gtggcggttg	tcgtgggtgcc	ttttttggtg	cccgtcacaa	60
actcaatttt	tgctcgctcc	tttttggcct	tttccaattt	gtccatctca	attttctggg	120
ccttggctaa	tgccatcatag	taggagtcct	cagaccagcc	atggggatca	aacatatcct	180
ttgggtagtt	ggtgcccaagc	tcgtcaatgg	cacagaatgg	atcagcttct	cgtaaatacta	240
gggttccgaa	attcttttctt	ccttttgata	atgtagttca	tatccattcc	ctccttttatc	300
t						301

<210> 238

<211> 301

<212> DNA

<213> Homo sapien

<400> 238

gggcaggttt	tttttttttt	ttttttgatg	gtgcagaccc	ttgctttatt	tgtctgactt	60
gttcacagtt	cagccccctg	ctcagaaaac	caacgggcca	gctaaggaga	ggaggaggca	120
ccttgagact	tccggagtcg	aggtctctcca	gggttcccca	gcccatcaat	cattttctgc	180
acccctgcc	tgggaagcag	ctccctgggg	ggtgggaatg	ggtgactaga	agggatttca	240
gtgtgggacc	cagggtctgt	tcttcacagt	aggaggtgga	agggatgact	aattttcttta	300
t						301

<210> 239

<211> 239

<212> DNA

<213> Homo sapien

<400> 239

ataagcagct agggaattct ttatttagta atgtcctaac ataaaagtgc acataactgc	60
ttctgtcaaa ccatgatact gagctttgtg acaaccaga aataactaag agaaggcaaa	120
cataatacct tagagatcaa gaaacattta cacagttcaa ctgtttaaaa atagctcaac	180
attcagccag tgagtagagt gtgaatgcc gcatacacag tatacaggtc cttcaggga	239

<210> 240

<211> 300

<212> DNA

<213> Homo sapien

<400> 240

ggtcctaattg aagcagcagc ttccacattt taacgcaggt ttacggtgat actgtccttt	60
gggatctgcc ctccagtga acccttttaag gaagaagtgg gcccaagcta agttccacat	120
gctgggtgag ccagatgact tctgttcctt ggtcactttc ttcaatgggg cgaatggggg	180
ctgccaggtt tttaaaatca tgcttcatct tgaagcacac ggtcacttca ccctcctcac	240
gctgtgggtg tactttgatg aaaataccca ctttggtggc ctttctgaag ctataatgtc	300

<210> 241

<211> 301

<212> DNA

<213> Homo sapien

<400> 241

gaggtctggt gctgaggtct ctgggctagg aagaggagtt ctgtggagct ggaagccaga	60
cctctttgga ggaaactcca gcagctatgt tgggtgtctct gagggaatgc aacaaggctg	120
ctcctccatg tattggaaaa ctgcaaactg gactcaactg gaaggaagtg ctgctgccag	180
tgtaagaac cagcctgagg tgacagaaac ggaagcaaac aggaacagcc agtcttttct	240
tcctcctcct gtcatacggg ctctctcaag catcctttgt tgtcaggggc ctaaaaggga	300
g	301

<210> 242

<211> 301

<212> DNA

<213> Homo sapien

<400> 242

ccgaggtcct gggatgcaac caatcactct gtttcacgtg acttttatca ccatacaatt	60
tgtggcattt cctcattttc tacattgtag aatcaagagt gtaaataaat gtatatcgat	120
gtcttcaaga atatatcatt cctttttcac tagaaccat tcaaaatata agtcaagaat	180
cttaatatca acaaatatat caagcaaact ggaaggcaga ataactacca taatttagta	240
taagtaccca aagttttata aatcaaaagc cctaatagata accattttta gaattcaatc	300
a	301

<210> 243

<211> 301

<212> DNA

<213> Homo sapien

<400> 243

```

aggtaagtcc cagtttgaag ctcaaaagat ctggtatgag cataggctca tcgacgacat      60
ggtggcccaa gctatgaaat cagagggagg cttcatcttg gcctgtaaaa actatgatgg      120
tgacgtgcag tcggactctg tggcccaagg gtatggctct ctcggcata gaaccagcgt      180
gctggtttgt ccagatggca agacagtaga agcagaggct gccacggga ctgtaacccg      240
tcactaccgc atgttccaga aaggacagga gacgtccacc aatcccattg cttccatttt      300
t                                                                                   301

```

```

<210> 244
<211> 300
<212> DNA
<213> Homo sapien

```

```

<400> 244
gctggtttgc aagaatgaaa tgaatgattc tacagctagg acttaacctt gaaatggaaa      60
gtcatgcaat cccatttgca ggatctgtct gtgcacatgc ctctgtagag agcagcattc      120
ccagggacct tggaaacagt tgacactgta aggtgcttgc tccccagac acatcctaaa      180
aggtgttgta atggtgaaaa cgtcttctct ctttattgcc ctttcttatt tatgtgaaca      240
actgtttgtc ttttgtgtat cttttttaaa ctgtaaagtt caattgtgaa aatgaatatc      300

```

```

<210> 245
<211> 301
<212> DNA
<213> Homo sapien

```

```

<400> 245
gtctgagtat ttaaaatggt attgaaatta tccccacca atgttagaaa agaaagaggt      60
tatatactta gataaaaaat gaggtgaatt actatccatt gaaatcatgc tcttagaatt      120
aaggccagga gatattgtca ttaatgtara cttcaggaca ctagagtata gcagccctat      180
gttttcaaag agcagagatg caattaaata ttgttttagca tcaaaaaggc cactcaatac      240
agctaataaa atgaaagacc taatttctaa agcaattctt tataatttac aaagttttaa      300
g                                                                                   301

```

```

<210> 246
<211> 301
<212> DNA
<213> Homo sapien

```

```

<400> 246
ggtctgtcct acaatgcctg cttcttgaaa gaagtcggca ctttctagaa tagctaaata      60
acctgggctt atttttaaaga actatttgta gctcagattg gttttcctat ggctaaaata      120
agtgttctt gtgaaaatta aataaaacag ttaattcaaa gccttgatat atgttaccac      180
taacaatcat actaaatata ttttgaagta caaagtttga catgctctaa agtgacaacc      240
caaatgtgtc ttacaaaaca cgttcttaac aaggtatgct ttacactacc aatgcagaaa      300
c                                                                                   301

```

```

<210> 247
<211> 301
<212> DNA
<213> Homo sapien

```

```

<400> 247

```

```

aggtcctttg gcagggctca tggatcagag ctcaaactgg agggaaaggc atttcgggta      60
gcctaagagg gcgactggcg gcagcacaac caaggaaggc aagggtgttt cccccacgct    120
gtgtcctgtg ttcagggtgc acacacaatc ctcatgggaa caggatcacc catgcgctgc    180
ccttgatgat caaggttggg gcttaagtgg attaagggag gcaagttctg ggttccttgc    240
cttttcaaac catgaagtca ggctctgtat ccctcctttt cctaactgat attctaacta    300
a                                                                 301

```

```

<210> 248
<211> 301
<212> DNA
<213> Homo sapien

```

```

<400> 248
aggtccttgg agatgccatt tcagccgaag gactcttctw ttcggaagta caccctcact      60
attaggaaga ttcttagggg taatttttct gaggaaggag aactagccaa cttagaatt      120
acaggaagaa agtggtttgg aagacagcca aagaaataaa agcagattaa attgtatcag    180
gtacattcca gcctgttggc aactccataa aaacatttca gattttaatc ccgaatttag    240
ctaattgagac tggatttttg ttttttatgt tgtgtgtcgc agagctaaaa actcagttcc    300
c                                                                 301

```

```

<210> 249
<211> 301
<212> DNA
<213> Homo sapien

```

```

<400> 249
gtccagagga agcacctggg gctgaactag gcttgccctg ctgtgaactt gcacttggag      60
ccctgacgct gctgttctcc ccgaaaaacc cgaccgacct ccgcgatctc cgtcccgcgc    120
ccagggagac acagcagtga ctcagagctg gtcgcacact gtgcctccct cctcaccgcc    180
catcgtaatg aattattttg aaaattaatt ccaccatcct ttcagattct ggatggaaag    240
actgaatctt tgactcagaa ttgtttgctg aaaagaatga tgtgactttc ttagtcattt    300
a                                                                 301

```

```

<210> 250
<211> 301
<212> DNA
<213> Homo sapien

```

```

<400> 250
ggctctgtgac aaggacttgc aggctgtggg aggcaagtga cccttaacac tacacttctc      60
cttatcttta ttggcttgat aaacataatt atttctaaca ctagcttatt tccagttgcc    120
cataagcaca tcagtacttt tctctggctg gaatagtaaa ctaaagtatg gtacatctac    180
ctaaaagact actatgtgga ataatacata ctaatgaagt attacatgat ttaaagacta    240
caataaaacc aaacatgctt ataacattaa gaaaaacaat aaagatacat gattgaaacc    300
a                                                                 301

```

```

<210> 251
<211> 301
<212> DNA
<213> Homo sapien

```

<400> 251

```

gccgaggtcc tacatattggc ccagtttccc cctgcatacct ctccagggcc cctgcctcat      60
agacaacctc atagagcata ggagaactgg ttgccctggg ggcaggggga ctgtctggat      120
ggcaggggtc ctcaaaaatg ccactgtcac tgccaggaaa tgcttctgag cagtacacct      180
cattgggatc aatgaaaagc ttcaagaaat cttcaggctc actctcttga aggcccggaa      240
cctctggagg ggggcagtgg aatcccagct ccaggacgga tcctgtcgaa aagatatcct      300
c                                                                                   301

```

<210> 252

<211> 301

<212> DNA

<213> Homo sapien

<400> 252

```

gcaaccaatc actctgtttc acgtgacttt tatcaccata caatttgtgg catttcctca      60
ttttctacat tgtagaatca agagtgtaaa taaatgtata tcgatgtctt caagaatata      120
tcattccttt ttcactagga acccattcaa aatataagtc aagaatctta atatcaacaa      180
atatatcaag caaactggaa ggcagaataa ctaccataat ttagtataag tacccaaagt      240
tttataaatc aaaagcccta atgataacca tttttagaat tcaatcatca ctgtagaatc      300
a                                                                                   301

```

<210> 253

<211> 301

<212> DNA

<213> Homo sapien

<400> 253

```

ttccctaaga agatgttatt ttgttgggtt ttgttcccc tccatctcga ttctcgtacc      60
caactaaaaa aaaaaaataa agaaaaaatg tgctgcgttc tgaaaaataa ctccttagct      120
tggtctgatt gttttcagac cttaaaatat aaacttgttt cacaagcttt aatccatgtg      180
gatttttttt cttagagaac cacaaaacat aaaaggagca agtcggactg aatacctgtt      240
tccatagtgc ccacagggta ttccctcacat tttctccata ggaaaatgct ttttcccaag      300
g                                                                                   301

```

<210> 254

<211> 301

<212> DNA

<213> Homo sapien

<400> 254

```

cgctgcgcct ttcccttggg ggagggggcaa ggccagaggg ggtccaagtg cagcacgagg      60
aacttgacca attcccttga agcgggtggg ttaaaccctg taaatgggaa caaatcccc      120
ccaaatctct tcattctacc ctggtggact cctgactgta gaattttttg gttgaaacaa      180
gaaaaaaata aagcttttga cttttcaagg ttgcttaaca ggtactgaaa gactggcctc      240
acttaaaactg agccaggaaa agctgcagat ttattaatgg gtgtgttagt gtgcagtgcc      300
t                                                                                   301

```

<210> 255

<211> 302

<212> DNA

<213> Homo sapien

<400> 255

```

agcttttttt tttttttttt tttttttttt ttcattaaaa aatagtgtct tttattataa      60
attactgaaa tgtttctttt ctgaatataa atataaatat gtgcaaagtt tgacttggat      120
tgggattttg ttgagttctt caagcatctc ctaataccct caagggcctg agtagggggg      180
aggaaaaagg actggagggtg gaatctttat aaaaaacaag agtgattgag gcagattgta      240
aacattatta aaaaacaaga aacaaacaaa aaaatagaga aaaaaaccac cccaacacac      300
aa                                                                    302

```

<210> 256

<211> 301

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(301)

<223> n = A,T,C or G

<400> 256

```

gttccagaaa acattgaagg tggcttccca aagtctaact agggataccc cctctagcct      60
aggaccctcc tccccacacc tcaatccacc aaaccatcca taatgcaccc agataggccc      120
acccccaaaa gcctggacac cttgagcaca cagttatgac caggacagac tcctctctat      180
aggcaaatag ctgctggcaa actggcatta cctggtttgt ggggatgggg gggcaagtgt      240
gtggcctctc ggctgggtta gcaagaacat tcagggtagg cctaagttan tcgtgttagt      300
t                                                                    301

```

<210> 257

<211> 301

<212> DNA

<213> Homo sapien

<400> 257

```

gttgtggagg aactctggct tgctcattaa gtcctactga ttttcactat cccctgaatt      60
tccccactta tttttgtctt tcaatatcgc aggccttaga agaggtctac ctgcctccag      120
tcttacctag tccagtctac cccctggagt tagaatggcc atcctgaagt gaaaagtaat      180
gtcacattac tcccttcagt gatctcttgt agaagtgcc atccctgaat gccaccaaga      240
tcttaatctt cacatcttta atcttatctc tttgactcct ctttacaccg gagaaggctc      300
c                                                                    301

```

<210> 258

<211> 301

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(301)

<223> n = A,T,C or G

<400> 258

<210> 262
 <211> 301
 <212> DNA
 <213> Homo sapien

<400> 262
 gaggagagcc tgttacagca tttgtaagca cagaatactc caggagtatt tgtaattgtc 60
 tgtgagcttc ttgccgcaag tctctcagaa atttaaaaag atgcaaacc ctgagtcacc 120
 cctagacttc ctaaaccaga tcctctgggg ctggaacctg gcaactctgca tttgtaatga 180
 gggctttctg gtgcacacct aattttgtgc atctttgccc taaatcctgg attagtgtcc 240
 catcattacc cccacattat aatgggatag attcagagca gatactctcc agcaaagaat 300
 c 301

<210> 263
 <211> 301
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(301)
 <223> n = A,T,C or G

<400> 263
 tttagcttgt ggtaaatgac tcacaaaact gattttaaaa tcaagttaat gtgaattttg 60
 aaaattacta cttaatccta attcacaata acaatggcat taagggttga cttgagttgg 120
 ttcttagtat tatttatggg aaataggctc ttaccacttg caaataactg gccacatcat 180
 taatgactga cttcccagta aggctctcta aggggtaagt angaggatcc acaggatttg 240
 agatgctaag gccccagaga tcgtttgatc caaccctctt attttcagag gggaaaatgg 300
 g 301

<210> 264
 <211> 301
 <212> DNA
 <213> Homo sapien

<400> 264
 aaagacgtta aaccactcta ctaccacttg tggaactctc aaagggtaaa tgacaaascc 60
 aatgaatgac tctaaaaaca atattttacat ttaatggttt gtagacaata aaaaaacaag 120
 gtggatagat ctagaattgt aacattttta gaaaaccata scatttgaca gatgagaaag 180
 ctcaattata gatgcaaagt tataactaaa ctactatagt agtaaagaaa tacatttcac 240
 acccttcata taaattcact atcttggtt gaggcactcc ataaaatgta tcacgtgcat 300
 a 301

<210> 265
 <211> 301
 <212> DNA
 <213> Homo sapien

<400> 265
 tgcccaagtt atgtgtaagt gtatccgcac ccagaggtaa aactacactg tcattcttgt 60


```

taacaatata cactagctat ctttttaact gtccatcatt agcaccaatg aagattcaat      60
aaaattacct ttattcacac atctcaaaac aattctgcaa attcttagtg aagtttaact     120
atagtcacag accttaaata ttcacattgt tttctatgtc tactgaaaat aagttcacta     180
cttttctgga tattctttac aaaatcttat taaaattcct ggtattatca cccccaatta     240
tacagtagca caaccacctt atgtagtttt tacatgatag ctctgtagaa gtttcacatc     300
t                                                                 301

```

```

<210> 270
<211> 301
<212> DNA
<213> Homo sapien

```

```

<400> 270
cattgaagag cttttgcgaa acatcagaac acaagtgcct ataaaattaa ttaagcctta      60
cacaagaata catattcctt ttatttctaa ggagttaaac atagatgtag ctgatgtgga     120
gagcttgctg gtgcagtgc aattggataa cactattcat ggccgaattg atcaagtcaa     180
ccaactcctt gaactggatc atcagaagaa ggggtggtgca cgatatactg cactagataa     240
tggaccaacc aactaaattc tctcaccagg ctgtatcagt aaactggcct aacagaaaac     300
a                                                                 301

```

```

<210> 271
<211> 301
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(301)
<223> n = A,T,C or G

```

```

<400> 271
aaaaggttct cataagatta acaatttaaa taaatatttg atagaacatt ctttctcatt      60
tttatagctc atcttttaggg ttgatattca gttcatgcct cccttgctgt tcttgatcca     120
gaattgcaat cacttcatca gcctgtattc gctccaattc tctataaagt ggggtccaagg     180
tgaaccacag agccacagca cacctctttc ccttggtgac tgccttcacc ccatganggt     240
tctctcctcc agatganaac tgatcatgog cccacatttt gggttttata gaagcagtca     300
c                                                                 301

```

```

<210> 272
<211> 301
<212> DNA
<213> Homo sapien

```

```

<400> 272
taaattgcta agccacagat aacaccaatc aaatggaaca aatcactgtc ttcaaattgtc      60
ttatcagaaa accaaatgag cctggaatct tcataatacc taaacatgcc gtatttagga     120
tccaataatt ccctcatgat gagcaagaaa aattctttgc gcacccctcc tgcattccaca     180
gcatcttctc caacaaatat aaccttgagt ggcttcttgt aatctatgtt ctttgttttc     240
ctaaggactt ccattgcac tctacaata ttttctctac gcaccactag aattaagcag     300
g                                                                 301

```


<400> 286

```
<210> 287
<211> 301
<212> DNA
<213> Homo sapien
```

<400> 287

```
<210> 288
<211> 301
<212> DNA
<213> Homo sapien
```

<400> 288

```
<210> 289
<211> 301
<212> DNA
<213> Homo sapien
```

<220>

```
<221> misc_feature
<222> (1)...(301)
<223> n = A,T,C or G
```

<400> 289

ggtacactgt ttccatgtta tgtttctaca cattgctacc tcagtgtccc tggaaactta 60

```

gcttttgatg tctccaagta gtccaccttc atttaactct ttgaaactgt atcatctttg      120
ccaagtaaga gtggtggcct atttcagctg ctttgacaaa atgactggct cctgacttaa      180
cgttctataa atgaatgtgc tgaagcaaag tgcccatggt ggcggcgaan aagagaaaga      240
tgtgttttgt tttggactct ctgtggtccc ttccaatgct gtgggtttcc aaccagnnga      300
a                                                                           301

```

```

<210> 290
<211> 301
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(301)
<223> n = A,T,C or G

```

```

<400> 290
acactgagct cttcttgata aatatacaga atgcttggca tatacaagat tctatactac      60
tgactgatct gttcatttct ctcacagctc ttaccccaa aagcttttcc accctaagtg      120
ttctgacctc cttttctaata cacagtaggg atagaggcag anccacctac aatgaacatg      180
gagttctatc aagaggcaga aacagcacag aatcccagtt ttaccattcg ctagcagtgc      240
tgccttgaac aaaaacattt ctccatgtct cattttcttc atgcctcaag taacagtgag      300
a                                                                           301

```

```

<210> 291
<211> 301
<212> DNA
<213> Homo sapien

```

```

<400> 291
caggtaccaa tttcttctat cctagaaaca tttcatttta tgttgttgaa acataacaac      60
tatatcagct agattttttt tctatgcttt acctgctatg gaaaatttga cacattctgc      120
tttactcttt tgtttatagg tgaatcacia aatgtatttt tatgtattct gtagttcaat      180
agccatggct gtttacttca tttaatttat ttagcataaa gacattatga aaaggcctaa      240
acatgagctt cacttcccca ctaactaatt agcatctggt atttcttaac cgtaatgcct      300
a                                                                           301

```

```

<210> 292
<211> 301
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(301)
<223> n = A,T,C or G

```

```

<400> 292
accttttagt agtaatgtct aataataaat aagaaatcaa ttttataagg tccatatagc      60
tgtattaaat aatttttaag tttaaaagat aaaataccat cattttaaat gttggtattc      120
aaaaccaaag natataaccg aaaggaaaaa cagatgagac ataaaaatgat ttgcnagatg      180

```

```

ggaaatatag tasttyatga atgttnatta aattccagtt ataatagtgg ctacacactc 240
tcactacaca cacagacccc acagtcctat atgccacaaa cacatttcca taacttgaaa 300
a 301

```

```

<210> 293
<211> 301
<212> DNA
<213> Homo sapien

```

```

<400> 293
ggtaccaagt gctggtgcca gcctgttacc tgttctcact gaaaagtctg gctaattgctc 60
ttgtgtagtc acttctgatt ctgacaatca atcaatcaat ggcttagagc actgactggt 120
aacacaaacg tcactagcaa agtagcaaca gctttaagtc taaatacaaa gctgttctgt 180
gtgagaattt tttaaaaggc tacttgtata ataacccttg tcattttttaa tgtacctcgg 240
ccgcgaccac gctaagccga attctgcaga tatccatcac actggcgggc gctcgagcat 300
g 301

```

```

<210> 294
<211> 301
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(301)
<223> n = A,T,C or G

```

```

<400> 294
tgacccataa caatatacac tagctatctt tttaaactgtc catcattagc accaatgaag 60
attcaataaa attaccttta ttcacacatc tcaaaacaat tctgcaaatt cttagtgaag 120
tttaactata gtcacaganc ttaaatattc acattgtttt ctatgtctac tgaaaataag 180
ttcactactt ttctgggata ttctttacaa aatcttatta aaattcctgg tattatcacc 240
ccaattata cagtagcaca accaccttat gtagttttta catgatagct ctgtagaggt 300
t 301

```

```

<210> 295
<211> 305
<212> DNA
<213> Homo sapien

```

```

<400> 295
gtactctttc tctccctcc tctgaattta attctttcaa cttgcaattt gcaaggatta 60
cacatttcac tgtgatgtat attgtgttgc aaaaaaaaaa gtgtctttgt ttaaaattac 120
ttggtttgat aatccatctt gctttttccc cattggaact agtcattaac ccactctctga 180
actggtagaa aaacrtctga agagctagtc tatcagcatc tgacaggtga attggatggt 240
tctcagaacc atttcaccca gacagcctgt ttctatcctg ttttaataaat tagtttgggt 300
tctct 305

```

```

<210> 296
<211> 301
<212> DNA

```

<213> Homo sapien

<400> 296

```

aggtactatg ggaagctgct aaaataatat ttgatagtaa aagtatgtaa tgtgctatct      60
cacctagtag taaactaaaa ataaactgaa actttatgga atctgaagtt attttccttg      120
attaaataga attaataaac caatatgagg aaacatgaaa ccatgcaatc tactatcaac      180
tttgaaaaag tgattgaacg aaccacttag ctttcagatg atgaacactg ataagtcatt      240
tgtcattact ataaatttta aaatctgtta ataagatggc ctatagggag gaaaaagggg      300
c                                                                                   301

```

<210> 297

<211> 300

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(300)

<223> n = A,T,C or G

<400> 297

```

actgagtttt aactggacgc caagcaggca aggctggaag gttttgctct ctttgtgcta      60
aagggttttg aaaccttgaa ggagaatcat ttgacaaga agtacttaag agtctagaga      120
acaaagangt gaaccagctg aaagctctcg ggggaanctt acatgtgttg ttaggcctgt      180
tccatcattg ggagtgcact ggccatccct caaaatttgt ctgggctggc ctgagtggtc      240
accgcacctc ggccgcgacc acgctaagcc gaattctgca gatatccatc acactggcgg      300

```

<210> 298

<211> 301

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(301)

<223> n = A,T,C or G

<400> 298

```

tatggggttt gtcacccaaa agctgatgct gagaaaggcc tccttggggc cctccccgcg      60
ggcatctgag agacctggtg ttccagtgtt tctggaaatg ggtcccagtg ccgccggctg      120
tgaagctctc agatcaatca cgggaagggc ctggcggttg tggccacctg gaaccaccct      180
gtcctgtctg tttacatttc actaycaggt tttctctggg cattacnatt tgttccccta      240
caacagtgac ctgtgcattc tgctgtggcc tgctgtgtct gcaggtggct ctcagcgagg      300
t                                                                                   301

```

<210> 299

<211> 301

<212> DNA

<213> Homo sapien

<400> 299

```

gttttgagac ggagtttcac tcttggtgcc cagactggac tgcaatggca ggggtctctgc      60
tactgcacc ctctgcctcc caggttcgag caattctcct gcctcagcct cccaggttagc      120
tggtgattgca ggctcacgcc accataccca gctaattttt ttgtattttt agtagagacg      180
gagtttcgcc atgttgggca gctggtctca aactcctgac ctcaagcgac ctgcctgcct      240
cggcctccca aagtgctgga attataggca tgagtcaaca cgcccagcct aaagatattt      300
t                                                                 301

```

<210> 300

<211> 301

<212> DNA

<213> Homo sapien

<400> 300

```

attcagtttt atttgctgcc ccagtatctg taaccaggag tgccacaaaa tcttgccaga      60
tatgtcccac acccactggg aaaggctccc acctggctac ttctctatc agctgggtca      120
gctgcattcc acaaggttct cagcctaatt agtttcaacta cctgccagtc tcaaaactta      180
gtaaagcaag accatgacat tccccacgg aaatcagagt ttgccccacc gtcttggttac      240
tataaagcct gcctctaaca gtccttgctt cttcacacca atcccagcgc catcccccat      300
g                                                                 301

```

<210> 301

<211> 301

<212> DNA

<213> Homo sapien

<400> 301

```

ttaaattttt gagaggataa aaaggacaaa taatctagaa atgtgtcttc ttcagtctgc      60
agaggacccc aggtctccaa gcaaccacat ggtcaagggc atgaataatt aaaagttggt      120
gggaactcac aaagaccctc agagctgaga ccccacaac agtgggagct caciaagacc      180
ctcagagctg agacaccac aacagtggga gctcacaag accctcagag ctgagacacc      240
cacaacagca cctcgttcag ctgccacatg tgtgaataag gatgcaatgt ccagaagtgt      300
t                                                                 301

```

<210> 302

<211> 301

<212> DNA

<213> Homo sapien

<400> 302

```

aggtacacat ttagcttggt gtaaatagact cacaaaactg attttaaaat caagttaatg      60
tgaattttga aaattactac ttaatcctaa ttcacaataa caatggcatt aaggtttgac      120
ttgagttggt tcttagtatt atttatggta aataggctct taccacttgc aaataactgg      180
ccacatcatt aatgactgac ttcccagtaa ggctctctaa ggggtaagta ggaggatcca      240
caggatttga gatgctaagg cccagagat cgtttgatcc aaccctctta ttttcagagg      300
g                                                                 301

```

<210> 303

<211> 301

<212> DNA

<213> Homo sapien

<400> 303
 aggtaccaac tgtggaaata ggtagaggat cattttttct tccatatca actaagttgt 60
 atattgtttt ttgacagttt aacacatctt cttctgtcag agattctttc acaatagcac 120
 tggctaattg aactaccgct tgcattgttaa aaatggtggt ttgtgaaatg atcataggcc 180
 agtaacgggt atgtttttct aactgatctt ttgctcgttc caaagggacc tcaagacttc 240
 catcgatttt atatctgggg tctagaaaag gagttaatct gttttccctc ataaattcac 300
 c 301

<210> 304
 <211> 301
 <212> DNA
 <213> Homo sapien

<400> 304
 acatggatgt tatttttcag actgtcaacc tgaatttgta tttgcttgac attgcctaatt 60
 tattagtttc agtttcagct taccactttt ttgtctgcaa catgcaraas agacagtgcc 120
 ctttttagtg tatcatatca ggaatcatct cacattgggt tgtgccatta ctggtgcagt 180
 gactttcagc cacttgggta aggtggagtt ggccatatgt ctccactgca aaattactga 240
 ttttcctttt gtaattaata agtgtgtgtg tgaagattct ttgagatgag gtatatactt 300
 c 301

<210> 305
 <211> 301
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(301)
 <223> n = A,T,C or G

<400> 305
 gangtacagc gtggtcaagg taacaagaag aaaaaaatgt gagtggcatc ctgggatgag 60
 cagggggaca gacctggaca gacacgttgt catttgctgc tgtgggtagg aaaatgggcg 120
 taaaggagga gaaacagata caaatctcc aactcagtat taaggtattc tcatgcctag 180
 aatattggta gaaacaagaa tacattcata tggcaaataa ctaaccatgg tggaacaaaa 240
 ttctgggatt taagttggat accaangaaa ttgtattaaa agagctgttc atggaataag 300
 a 301

<210> 306
 <211> 8
 <212> PRT
 <213> Homo sapien

<400> 306
 Val Leu Gly Trp Val Ala Glu Leu
 1 5

<210> 307
 <211> 637
 <212> DNA

<213> Homo sapien

<400> 307

```
acaggggratg aagggaaagg gagaggatga ggaagccccc ctgggggattt ggtttggtcc      60
ttgtgatcag gtggtctatg gggcttatcc ctacaaagaa gaatccagaa atagggggcac      120
attgaggaat gatacttgag cccaaagagc attcaatcat tgttttattt gccttmtttt      180
cacaccattg gtgagggagg gattaccacc ctgggggttat gaagatgggtt gaacacccca      240
cacatagcac cggagatatg agatcaacag tttcttagcc atagagattc acagcccaga      300
gcaggaggac gcttgcacac catgcaggat gacatggggg atgcgctcgg gattgggtgtg      360
aagaagcaag gactgttaga ggcaggcttt atagtaacaa gacgggtgggg caaactctga      420
tttccgtggg ggaatgtcat ggtcttgctt tactaagttt tgagactggc aggtagtga      480
actcattagg ctgagaacct tgtggaatgc acttgacca sctgatagag gaagtagcca      540
ggtagggagcc tttccagtg ggtgtgggac atatctggca agattttgtg gcactcctgg      600
ttacagatac tggggcagca aataaaaactg aatcttg      637
```

<210> 308

<211> 647

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(647)

<223> n = A,T,C or G

<400> 308

```
acgattttca ttatcatgta aatcgggtca ctcaaggggc caaccacagc tgggagccac      60
tgctcagggg aaggttcata tgggactttc tactgoccaa ggttctatac aggatataaa      120
ggngcctcac agtatagatc tggtagcaaa gaagaagaaa caaactctga tctctttctg      180
ccaccctctc gacccttttg aactcctctg acccttttaga acaagcctac ctaatatctg      240
ctagagaaaa gaccaacaac ggcctcaaag gatctcttac catgaaggtc tcagctaatt      300
cttggctaag atgtgggttc cacattaggt tctgaatatg gggggaaggg tcaatttgct      360
cattttgtgt gtggataaag tcaggatgcc cagggggccag agcagggggc tgcttgcttt      420
gggaacaatg gctgagcata taaccatagg ttatggggaa caaaacaaca tcaaagtcac      480
tgtatcaatt gccatgaaga cttgagggac ctgaatctac cgattcatct taaggcagca      540
ggaccagttt gagtggcaac aatgcagcag cagaatcaat ggaaacaaca gaatgattgc      600
aatgtccttt tttttctcct gcttctgact tgataaaagg ggaccgt      647
```

<210> 309

<211> 460

<212> DNA

<213> Homo sapien

<400> 309

```
actttatagt ttaggctgga cattggaaaa aaaaaaaagc cagaacaaca tgtgatagat      60
aatatgattg gctgcacact tccagactga tgaatgatga acgtgatgga ctattgtatg      120
gagcacatct tcagcaagag ggggaaatac tcatcatttt tggccagcag ttgtttgatc      180
accaaaccatc atgccagaat actcagcaaa ccttcttagc tcttgagaag tcaaagtcag      240
ggggaattta ttcttgcaa ttttaatttg actccttatg tgagagcagc ggctaccag      300
ctggggtggt ggagcgaacc cgtcactagt ggacatgcag tggcagagct cctggtaacc      360
acctagagga atacacaggc acatgtgtga tgccaagcgt gacacctgta gcactcaaat      420
```

ttgtcttgtt tttgtctttc ggtgtgtaag attcttaagt

460

<210> 310

<211> 539

<212> DNA

<213> Homo sapien

<400> 310

acgggactta	tcaaataaag	ataggaaaag	aagaaaactc	aatattata	ggcagaaatg	60
ctaaaggttt	taaaatatgt	caggattgga	agaaggcatg	gataaagaac	aaagttcagt	120
taggaaagag	aaacacagaa	ggaagagaca	caataaaagt	cattatgtat	tctgtgagaa	180
gtcagacagt	aagattttgt	ggaaatgggt	tggtttgttg	tatggtatgt	attttagcaa	240
taatctttat	ggcagagaaa	gctaaaatcc	tttagcttgc	gtgaatgac	acttgctgaa	300
ttcctcaagg	taggcatgat	gaaggagggt	ttagaggaga	cacagacaca	atgaactgac	360
ctagatagaa	agccttagta	tactcagcta	ggaatagtga	ttctgagggc	acactgtgac	420
atgattatgt	cattacatgt	atggtagtga	tggggatgat	aggaaggaag	aacttatggc	480
atattttcac	ccccacaaaa	gtcagttaaa	tattgggaca	ctaaccatcc	aggtcaaga	539

<210> 311

<211> 526

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(526)

<223> n = A,T,C or G

<400> 311

caaatttgag	ccaatgacat	agaattttac	aatcaagaa	gcttattctg	gggccatttc	60
ttttgacgtt	ttctctaaac	tactaaagag	gcattaatga	tccataaatt	atattatcta	120
catttacagc	atttaaaatg	tggtcagcat	gaaatattag	ctacagggga	agctaaataa	180
attaaacatg	gaataaagat	ttgtccttaa	atataatcta	caagaagact	ttgatatttg	240
tttttcacaa	gtgaagcatt	cttataaagt	gtcataacct	ttttggggaa	actatgggaa	300
aaaatgggga	aactctgaag	ggttttaagt	atcttacctg	aagctacaga	ctccataacc	360
tctctttaca	gggagctcct	gcagccccta	cagaaatgag	tggctgagat	tcttgattgc	420
acagcaagag	cttctcatct	aaaccctttc	ccttttttagt	atctgtgtat	caagtataaa	480
agttctataa	actgtagtnt	acttatttta	atccccaag	cacagt		526

<210> 312

<211> 500

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(500)

<223> n = A,T,C or G

<400> 312

cctctctctc	cccaccccct	gactctagag	aactggggtt	tctcccagta	ctccagcaat	60
------------	------------	------------	------------	------------	------------	----


```

tcattttctga aagcagttga gccactttat tccaaagtac actgcagatg ttcaaactct 120
ccattttctct ttccttcca cctgccagtt ttgtgactc tcaacttgtc atgagtgtaa 180
gcattaagga cattatgctt cttcgattct gaagacaggc cctgctcatg gatgactctg 240
gcttcttagg aaaatatttt tcttccaaaa tcagtaggaa atctaaactt atccccctct 300
tgcagatgtc tagcagcttc agacatttgg ttaagaacct atgggaaaaa aaaaaatcct 360
tgctaagtgt gtttcctttg taaaccanga ttcttatttg nctggtatag aatatcagct 420
ctgaacgtgt ggtaaagatt tttgtgtttg aatataggag aaatcagttt gctgaaaagt 480
tagtcttaat tatctattgg
500

```

```

<210> 313
<211> 718
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(718)
<223> n = A,T,C or G

```

```

<400> 313
ggagatttgt gtggtttgca gccgaggag accaggaaga tctgcatggt gggaaggacc 60
tgatgataca gaggtgagaa ataagaaagg ctgctgactt taccatctga ggccacacat 120
ctgctgaaat ggagataatt aacatcacta gaaacagcaa gatgacaata taatgtctaa 180
gtagtgacat gtttttgcac atttccagcc cttttaaata tccacacaca caggaagcac 240
aaaaggaagc acagagatcc ctggggagaaa tgcctggccg ccatcttggg tcatcgatga 300
gcctcgccct gtgcctgntc ccgcttgtga gggaaggaca ttagaaaatg aattgatgtg 360
ttccttaaag gatggcagga aaacagatcc tgttgtggat atttatttga acgggattac 420
agatttgaaa tgaagtcaca aagtgagcat taccaatgag aggaaaacag acgagaaaat 480
cttgatgggt cacaagacat gcaacaaaca aaatggaata ctgtgatgac acgagcagcc 540
aactggggag gagataccac ggggcagagg tcaggattct ggccctgctg cctaactgtg 600
cgttatacca atcattttcta tttctaccct caaacaagct gtngaataatc tgacttacgg 660
ttcttntggc ccacattttc atnatccacc ccttcttttt aannttantc caaantgt 718

```

```

<210> 314
<211> 358
<212> DNA
<213> Homo sapien

```

```

<400> 314
gtttattttac attacagaaa aaacatcaag acaatgtata ctatttcaaa tatatccata 60
cataatcaaa tatagctgta gtacatgttt tcattgggtg agattaccac aaatgcaagg 120
caacatgtgt agatctcttg tcttattctt ttgtctataa tactgtattg tgtagtccaa 180
gctctcggtg gtccagccac tgtgaaacat gctcccttta gattaacctc gtggacgctc 240
ttgttgattt gctgaactgt agtgccctgt attttgcttc tgtctgtgaa ttctgttgct 300
tctggggcat ttccttgtga tgcagaggac caccacacag atgacagcaa tctgaatt 358

```

```

<210> 315
<211> 341
<212> DNA
<213> Homo sapien

```

<400> 315
taccacctcc ccgctggcac tgatgagcgc catcaccatg gtcaccagca ccatgaaggc 60
ataggtgatg atgaggacat ggaatgggcc cccaaggatg gtctgtccaa agaagcgagt 120
gacccccatt ctgaagatgt ctggaacctc taccagcagg atgatgatag ccccaatgac 180
agtcaccagc tccccgacca gccggatata gtcccttaggg gtcattgtagg cttcctgaag 240
tagcttctgc tgtaagaggg tggtgtcccg ggggctcgtg cggttattgg tcttgggctt 300
gagggggcgg tagatgcagc acatggtgaa gcagatgatg t 341

<210> 316
<211> 151
<212> DNA
<213> Homo sapien

<400> 316
agactgggca agactcttac gcccacact gcaatttggc cttgttgccg tatccattta 60
tgtgggcctt tctcgagttt ctgattataa acaccactgg agcgatgtgt tgactggact 120
cattcaggga gctctggttg caatattagt t 151

<210> 317
<211> 151
<212> DNA
<213> Homo sapien

<400> 317
agaactagtg gatcctaata aaataacctga aacatatatt ggcattttatc aatggctcaa 60
atcttcattt atctctggcc ttaaccctgg ctcttgaggc tgcggccagc agatcccagg 120
ccagggtctt gttcttgcca cacctgcttg a 151

<210> 318
<211> 151
<212> DNA
<213> Homo sapien

<400> 318
actggtggga ggcgctgttt agttggctgt tttcagaggg gtctttcgga gggacctcct 60
gctgcaggct ggagtgtctt tattcctggc gggagaccgc acattccact gctgaggctg 120
tgggggcggt ttatcaggca gtgataaaca t 151

<210> 319
<211> 151
<212> DNA
<213> Homo sapien

<400> 319
aactagtgga tccagagcta taggtacagt gtgatctcag ctttgcaaac acattttcta 60
catagatagt actaggtatt aatagatatg taaagaaaga aatcacacca ttaataatgg 120
taagattggg tttatgtgat tttagtgggt a 151

<210> 320
<211> 150
<212> DNA

<400> 320

$\langle 210 \rangle$ 321

<211> 151

<212> DNA

<213> Homo sapien

 $\langle 400 \rangle$ 321

<210> 322

<211> 151

<212> DNA

<213> Homo sapien

<220>

<221> misc feature

 $\langle 222 \rangle \quad (1) \dots (151)$

<223> n = A, T, C or G

<400> 322

<210> 323

<211> 151

<212> DNA

<213> Homo sapien

 $\langle 220 \rangle$

<221> misc feature

 $\langle 222 \rangle \quad (1) \dots (151)$

<223> n = A, T, C or G

<400> 323

<210> 324

<211> 461

<212> DNA

<213> Homo sapien


```

gcttcttagc tgaggaaaag cacctccacg ttttgatcaa caatgcagga gtgatgatgt 420
gtccgtactc gaagacagca gatggctttg agatgcacat aggagtcaac cacttgggtc 480
acttcctcct aacccatctg ctgctagaga aactaaagga atcagcccca tcaaggatag 540
taaattgtgtc ttccctcgca catcacctgg gaaggatcca cttccataac ctgcagggcg 600
agaaattcta caatgcaggc ctggcctact gtcacagcaa gctagccaac atcctcttca 660
cccaggaact ggcccggaga ctaaaaggct ctggcggttac gacgtattct gtacaccctg 720
gcacagtcca atctgaactg gttcggcact catctttcat gagatggatg tgggtggcttt 780
tctccttttt catcaagact cctcagcagg gagcccagac cagcctgcac tgtgccttaa 840
cagaaggctct tgagattcta agtgggaatc atttcagtga ctgtcatgtg gcatgggtct 900
ctgcccgaagc tcgtaatgag actatagcaa ggcggctgtg ggacgtcagt tgtgacctgc 960
tgggcctccc aatagactaa caggcagtg cagttggacc caagagaaga ctgcagcaga 1020
ctacacagta cttcttgtca aaatgattct ccttcaagg tttcaaaacc tttagcacia 1080
agagagcaaa acctccagc cttgcctgct tgggtgtccag ttaaaactca gtgtactgcc 1140
agattcgtct aaatgtctgt catgtccaga tttactttgc ttctgttact gccagagtta 1200
ctagagatat cataatagga taagaagacc ctcatatgac ctgcacagct cattttcctt 1260
ctgaaagaaa ctactaccta ggagaatcta agctatagca gggatgattt atgcaaattt 1320
gaactagctt ctttgttcac aattcagttc ctcccaacca accagtcttc acttcaagag 1380
ggccacactg caacctcagc ttaacatgaa taacaaagac tggctcagga gcagggcttg 1440
cccaggcatg gtggatcacc ggaggtcagt agttcaagac cagcctggcc aacatgggtga 1500
aaccccacct ctactaaaaa ttgtgtatat ctttgttgtt cttcctgttt atgtgtgcca 1560
agggagtatt ttcacaaagt tcaaaacagc cacaataatc agagatggag caaaccagt 1620
ccatccagtc tttatgcaaa tgaaatgctg caaagggag cagattctgt atatgttgg 1680
aactaccac caagagcaca tgggtagcag ggaagaagta aaaaaagaga aggagaatac 1740
tggaagataa tgcacaaaat gaagggacta gttaaggatt aactagccct ttaaggatta 1800
actagttaag gattaatagc aaaagayatt aaatatgcta acatagctat ggaggaattg 1860
agggcaagca cccaggactg atgaggtctt aacaaaaacc agtgtggcaa aaaaaaaaaa 1920
aaaaaaaaaa aaaaatccta aaaacaaaca acaaaaaaaa acaattcttc attcagaaaa 1980
attatcttag ggactgatat tggtaattat ggtcaattta ataataattt ggggcatttc 2040
cttacattgt cttgacaaga ttaaaatgtc tgtgccaaaa ttttgtattt tatttgga 2100
cttcttatca aaagtaatgc tgccaaagga agtctaagga attagtagtg tcccatcac 2160
ttgtttggag tgtgctattc taaaagattt tgatttctct gaatgacaat tatattttaa 2220
ctttgggtggg ggaaagagtt ataggaccac agtcttcact tctgatactt gtaaattaat 2280
cttttattgc acttgttttg accattaagc tatatgttta gaaatgggtca ttttacggaa 2340
aaattagaaa aattctgata atagtgcaga ataaatgaat taatgtttta cttaatttat 2400
attgaactgt caatgacaaa taaaaattct ttttgattat tttttgtttt catttaccag 2460
aataaaaacg taagaattaa aagtttgatt acaaaaaaaaa aaaaaaa 2507

```

<210> 333

<211> 3030

<212> DNA

<213> Homo sapien

<400> 333

```

gcaggcgact tgcgagctgg gagcgattta aaacgctttg gattcccccg gcctgggtgg 60
ggagagcgag ctgggtgccc cctagattcc ccgccccgc acctcatgag ccgacctctg 120
gctccatgga gcccggaat tatgccacct tggatggagc caaggatatc gaaggcttgc 180
tgggagcggg agggggggcg aatctggtcg cccactcccc tctgaccagc caccagcgg 240
cgcctacgct gatgcctgct gtcaactatg ccccttggga tctgccaggc tcggcgagc 300
cgccaaagca atgccacca tgccctgggg tgccccagg gacgtcccca gctcccgtgc 360
cttatggtta ctttgaggc ggggtactact cctgccgagt gtcccgagc tcgctgaaac 420
cctgtgccca ggcagccacc ctggcgcgct accccgcgga gactccacg gccggggaag 480

```

agtaccgccag ycgccccact gagtttgcc tctatccggg atatccggga acctaccagc 540
 ctatggccag ttacctggac gtgtctgtgg tgcagactct ggggtgctcct ggagaaccgc 600
 gacatgactc cctgttgcc gtggacagtt accagtcttg ggcctctcgt ggtggctgga 660
 acagccagat gtgttgccag ggagaacaga acccaccagg tcccttttgg aaggcagcat 720
 ttgcagactc cagcgggcag caccctcctg acgcctgcgc ctttcgtcgc ggccgcaaga 780
 aacgcattcc gtacagcaag gggcagttgc gggagctgga gcgggagtat gcgggctaaca 840
 agttcatcac caaggacaag aggcgcaaga tctcggcagc caccagcctc tcggagcgcc 900
 agattaccat ctggtttcag aaccgcccgg tcaaagagaa gaaggttctc gccaaaggta 960
 agaacagcgc taccctttaa gagatctcct tgccctgggtg ggaggagcga aagtgggggt 1020
 gtccctgggga gaccaggaac ctgccaaagc caggctgggg ccaaggactc tgctgagagg 1080
 cccctagaga caacaccctt cccaggccac tggctgctgg actgttcctc aggagcggcc 1140
 tgggtaccca gtatgtgcag ggagacggaa ccccatgtga cagccactc caccagggtt 1200
 cccaaagaac ctggcccagt cataatcatt catcctgaca gtggcaataa tcacgataac 1260
 cagtactagc tgccatgac gttagcctca tattttctat ctagagctct gtagagcact 1320
 ttagaaaccg ctttcatgaa ttgagctaat tatgaataaa tttggaaggc gatccctttg 1380
 caggaagct ttctctcaga ccccttcca ttacacctct caccctggta acagcaggaa 1440
 gactgaggag aggggaacgg gcagattcgt tgtgtggctg tgatgtccgt ttagcatttt 1500
 tctcagctga cagctgggta ggtggacaat tgtagaggct gtctcttcct cctccttgt 1560
 ccacccata ggggtgaccc actggtcttg gaagcaccca tccttaatac gatgattttt 1620
 ctgtcgtgtg aaaatgaagc cagcaggctg cccctagtca gtccctcctt ccagagaaaa 1680
 agagatttga gaaagtgcct gggtaattca ccattaattt cctcccccaa actctctgag 1740
 tcttccctta atatttctgg tggttctgac caaagcaggt catggtttgt tgagcatttg 1800
 ggatccagc gaagtagatg tttgtagcct tgcatactta gcccttccca ggcacaaacg 1860
 gagtggcaga gtggtgcca cctgttttc ccagtccacg tagacagatt cacagtgcgg 1920
 aattctggaa gctggagaca gacgggctct ttgcagagcc gggactctga gagggacatg 1980
 agggcctctg cctctgtgtt cattctctga tgtcctgtac ctgggctcag tgcccggtgg 2040
 gactcatctc ctggccgcgc agcaaagcca gcgggttcgt gctggtcctt cctgcacctt 2100
 aggtgggggg tggggggcct gccggcgcat tctccacgat tgagcgcaca ggcctgaagt 2160
 ctggacaacc cgcagaaccg aagctccgag cagcgggtcg gtggcgagta gtggggtcgg 2220
 tggcgagcag ttggtgggtg gccgcggccg ccactacctc gaggacattt cctcccggga 2280
 gccagctctc ctagaaaccc cgcggcgccg gccgcagcca agtgtttatg gccgcgggtc 2340
 ggggtgggatc ctagccctgt ctctctcct ggggaaggagt gaggggtggga cgtgacttag 2400
 acacctacaa atctatttac caaagaggag cccgggactg agggaaaagg ccaaagagtg 2460
 tgagtgcagc cggactgggg gttcaggggg agaggacgag gaggaggaag atgaggtcga 2520
 tttcctgatt taaaaaatcg tccaagcccc gtggtccagc ttaaggtcct cggttacatg 2580
 cgccgctcag agcaggtcac tttctgcctt ccacgtcctc cttcaaggaa gccccatgtg 2640
 ggtagctttc aatatcgag gtctcttact ctctgcctct ataagctcaa accaccaac 2700
 gatcgggcaa gtaaaccccc tccctcgccg acttcggaac tggcgagagt tcagcgaga 2760
 tgggcctgtg gggagggggc aagatagatg agggggagcg gcatggtgcg gggtgacccc 2820
 ttggagagag gaaaaaggcc acaagagggg ctgccaccgc cactaacgga gatggcctg 2880
 gtagagacct ttgggggtct ggaacctctg gactccccat gctctaactc ccacactctg 2940
 ctatcagaaa cttaaacttg aggattttct ctgtttttca ctcgcaataa aytcagagca 3000
 aacaaaaaaa aaaaaaaaaa aaaactcgag 3030

<210> 334

<211> 2417

<212> DNA

<213> Homo sapien

<400> 334

ggcggccgct ctagagctag tgggatcccc cgggctgcac gaattcggca cgagtgagtt 60

ggagttttac ctgtattgtt ttaattttcaa caagcctgag gactagccac aaatgtaccc 120
 agttttacaaa tgaggaaaca ggtgcaaaaa ggttggttacc tgtcaaaggt cgtatgtggc 180
 agagccaaga tttgagccca gttatgtctg atgaacttag cctatgctct ttaaacttct 240
 gaatgctgac cattgaggat atctaaactt agatcaattg cattttccct ccaagactat 300
 ttacttatca atacaataat accaccttta ccaatctatt gttttgatac gagactcaaa 360
 tatgccagat atatgtaaaa gcaacctaca agctctctaa tcatgctcac ctaaaagatt 420
 cccgggatct aataggctca aagaaacttc ttctagaaat ataaaagaga aaattggatt 480
 atgcaaaaat tcattattaa tttttttcat ccatccttta attcagcaaa catttatctg 540
 ttgttgactt tatgcagtat ggctttttta ggattggggg acaggtgaag aacgggggtgc 600
 cagaatgcat cctcctacta atgaggctag tacacatttg catttttaaaa tgccctgtcc 660
 agctgggcat ggtggatcat gcctgtaatc tcaacattgg aaggccaagg caggaggatt 720
 gcttcagccc aggagttaa gaccagcctg ggcaacatag aaagacccca tctctcaatc 780
 aatcaatcaa tgccctgtct ttgaaaataa aactccttaa gaaagggtta atgggcaggg 840
 tgtggttagct catgcctata atacagcact ttgggaggct gaggcaggag gatcacttta 900
 gcccagaagt tcaagaccag cctgggcaac aagtgcaccc tcatctcaat tttttaataa 960
 aatgaatata tacataagga aagataaaaa gaaaagttaa atgaaagaat acagtataaa 1020
 acaaactctct tggacctaaa agtatTTTTg ttcaagccaa atattgtgaa tcacctctct 1080
 gtgttgagga tacagaatat ctaagcccag gaaactgagc agaaagtta tgtactaact 1140
 aatcaacccg aggcaaggca aaaatgagac taactaatca atccgaggca aggggcaa 1200
 tagacggaac ctgactctgg tctattaagc gacaactttc cctctgttgt atttttcttt 1260
 tattcaatgt aaaaggataa aaactctcta aaactaaaaa caatgtttgt caggagtta 1320
 aaaccatgac caactaatta tggggaatca taaaatatga ctgtatgaga tcttgatgg 1380
 ttacaaagtg taccactgt taatcacttt aaacattaat gaacttaaaa atgaatttac 1440
 ggagattgga atgtttcttt cctgttgtat tagttggctc aggttgccat aacaaaatac 1500
 cacagactgg gaggtttaag taacagaaat tcattttctca cagttctggg ggctggaagt 1560
 ccacgatcaa ggtgcaggaa aggcaggctt cattctgagg cccctctctt ggctcacatg 1620
 tggccaccct cccactgcgt gctcacatga cctctttgtg ctcttgaaa gaggggtgtg 1680
 gggacagagg gaaagagaag gagagggaac tctctgggtg ctctctttc aaggacccta 1740
 acctgggcca ctttgggcca ggcactgtgg ggtggggggg tgtggctgct ctgctctgag 1800
 tggccaagat aaagcaacag aaaaatgtcc aaagctgtgc agcaaagaca agccaccgaa 1860
 cagggatctg ctcatcagt tggggacctc caagtgggc accctggagg caagcccca 1920
 cagagcccat gcaaggtggc agcagcagaa gaagggaatt gtccctgtcc ttggcacatt 1980
 cctcaccgac ctggtgatgc tggacactgc gatgaatggg aatgtggatg agaatatgat 2040
 ggactcccag aaaaggagac ccagctgctc aggtggctgc aaatcattac agccttcac 2100
 ctggggagga actggggggc tggttctggg tcagagagca gccagtgag ggtgagagct 2160
 acagcctgtc ctgccagctg gatccccagt cccggctaac cagtaatcaa ggctgagcag 2220
 atcaggcttc ccggagctgg tcttgggaag ccagccctgg ggtgagttgg ctctgtctgt 2280
 ggtactgaga caatatgtc ataaattcaa tgcgcccttg tatccctttt tcttttttat 2340
 ctgtctacat ctataatcac tatgcatact agtctttgtt agtgtttcta ttcmaactaa 2400
 tagagatatg ttatact 2417

<210> 335

<211> 2984

<212> DNA

<213> Homo sapien

<400> 335

atccctcctt cccactctc ctttccagaa ggcacttggg gtcttatctg ttggactctg 60
 aaaacacttc aggcgcctt ccaaggcttc cccaaacccc taagcagccg cagaagcgt 120
 cccgagctgc cttctccac actcaggtga tcgagttgga gaggaagttc agccatcaga 180
 agtacctgtc ggccctgaa cgggccacc tggccaagaa cctcaagctc acggagacc 240

<400> 336
 Pro Ser Phe Pro Thr Leu Leu Ser Arg Arg His Leu Gly Ser Tyr Leu
 1 5 10 15
 Leu Asp Ser Glu Asn Thr Ser Gly Ala Leu Pro Arg Leu Pro Gln Thr
 20 25 30
 Pro Lys Gln Pro Gln Lys Arg Ser Arg Ala Ala Phe Ser His Thr Gln
 35 40 45
 Val Ile Glu Leu Glu Arg Lys Phe Ser His Gln Lys Tyr Leu Ser Ala
 50 55 60
 Pro Glu Arg Ala His Leu Ala Lys Asn Leu Lys Leu Thr Glu Thr Gln
 65 70 75 80
 Val Lys Ile Trp Phe Gln Asn Arg Arg Tyr Lys Thr Lys Arg Lys Gln
 85 90 95
 Leu Ser Ser Glu Leu Gly Asp Leu Glu Lys His Ser Ser Leu Pro Ala
 100 105 110
 Leu Lys Glu Glu Ala Phe Ser Arg Ala Ser Leu Val Ser Val Tyr Asn
 115 120 125
 Ser Tyr Pro Tyr Tyr Pro Tyr Leu Tyr Cys Val Gly Ser Trp Ser Pro
 130 135 140
 Ala Phe Trp
 145

<210> 337
 <211> 9
 <212> PRT
 <213> Homo sapien

<400> 337
 Ala Leu Thr Gly Phe Thr Phe Ser Ala
 1 5

<210> 338
 <211> 9
 <212> PRT
 <213> Homo sapien

<400> 338
 Leu Leu Ala Asn Asp Leu Met Leu Ile
 1 5

<210> 339
 <211> 318
 <212> PRT
 <213> Homo sapien

<400> 339
 Met Val Glu Leu Met Phe Pro Leu Leu Leu Leu Leu Leu Pro Phe Leu
 1 5 10 15
 Leu Tyr Met Ala Ala Pro Gln Ile Arg Lys Met Leu Ser Ser Gly Val

20 25 30
 Cys Thr Ser Thr Val Gln Leu Pro Gly Lys Val Val Val Val Thr Gly
 35 40 45
 Ala Asn Thr Gly Ile Gly Lys Glu Thr Ala Lys Glu Leu Ala Gln Arg
 50 55 60
 Gly Ala Arg Val Tyr Leu Ala Cys Arg Asp Val Glu Lys Gly Glu Leu
 65 70 75 80
 Val Ala Lys Glu Ile Gln Thr Thr Thr Gly Asn Gln Gln Val Leu Val
 85 90 95
 Arg Lys Leu Asp Leu Ser Asp Thr Lys Ser Ile Arg Ala Phe Ala Lys
 100 105 110
 Gly Phe Leu Ala Glu Glu Lys His Leu His Val Leu Ile Asn Asn Ala
 115 120 125
 Gly Val Met Met Cys Pro Tyr Ser Lys Thr Ala Asp Gly Phe Glu Met
 130 135 140
 His Ile Gly Val Asn His Leu Gly His Phe Leu Leu Thr His Leu Leu
 145 150 155 160
 Leu Glu Lys Leu Lys Glu Ser Ala Pro Ser Arg Ile Val Asn Val Ser
 165 170 175
 Ser Leu Ala His His Leu Gly Arg Ile His Phe His Asn Leu Gln Gly
 180 185 190
 Glu Lys Phe Tyr Asn Ala Gly Leu Ala Tyr Cys His Ser Lys Leu Ala
 195 200 205
 Asn Ile Leu Phe Thr Gln Glu Leu Ala Arg Arg Leu Lys Gly Ser Gly
 210 215 220
 Val Thr Thr Tyr Ser Val His Pro Gly Thr Val Gln Ser Glu Leu Val
 225 230 235 240
 Arg His Ser Ser Phe Met Arg Trp Met Trp Trp Leu Phe Ser Phe Phe
 245 250 255
 Ile Lys Thr Pro Gln Gln Gly Ala Gln Thr Ser Leu His Cys Ala Leu
 260 265 270
 Thr Glu Gly Leu Glu Ile Leu Ser Gly Asn His Phe Ser Asp Cys His
 275 280 285
 Val Ala Trp Val Ser Ala Gln Ala Arg Asn Glu Thr Ile Ala Arg Arg
 290 295 300
 Leu Trp Asp Val Ser Cys Asp Leu Leu Gly Leu Pro Ile Asp
 305 310 315

<210> 340

<211> 483

<212> DNA

<213> Homo sapien

<400> 340

```

gccgaggtct gccttcacac ggaggacacg agactgcttc ctcaagggtc cctgcctgcc      60
tggacactgg tgggagggcg tgtttagttg gctgttttca gaggggtctt tcggagggac      120
ctcctgctgc aggctggagt gtctttattc ctggcgggag accgcacatt ccaactgctga      180
ggttgtgggg gcggtttatc aggcagtgat aaacataaga tgtcatttcc ttgactccgg      240
ccttcaattt tctcttttggc tgacgacgga gtccgtggtg tcccgatgta actgaccct      300
gctccaaacg tgacatcact gatgctcttc tcgggggtgc tgatggcccg cttggtcacg      360
tgctcaatct cgccattcga ctcttgctcc aaactgtatg aagacacctg actgcacgtt      420

```

ttttctgggc ttccagaatt taaagtgaaa ggcagcactc ctaagctccg actccgatgc 480
ctg 483

<210> 341
<211> 344
<212> DNA
<213> Homo sapien

<400> 341
ctgctgctga gtcacagatt tcattataaa tagcctccct aaggaaaata cactgaatgc 60
tatttttact aaccatttcta tttttataga aatagctgag agttttctaaa ccaactctct 120
gctgccttac aagtattaaa tattttactt ctttccataa agagtagctc aaaatatgca 180
attaatttaa taattttctga tgatgggttt atctgcagta atatgtatat catctattag 240
aatttactta atgaaaaact gaagagaaca aaatttgtaa ccactagcac ttaagtactc 300
ctgattctta acattgtctt taatgaccac aagacaacca acag 344

<210> 342
<211> 592
<212> DNA
<213> Homo sapien

<400> 342
acagcaaaaa agaaactgag aagcccaaty tgctttcttg ttaacatcca cttatccaac 60
caatgtggaa acttcttata cttggttcca ttatgaagtt ggacaattgc tgctatcaca 120
cctggcaggt aaaccaatgc caagagagtg atggaaacca ttggcaagac tttgttgatg 180
accaggattg gaattttata aaaatattgt tgatgggaag ttgctaaagg gtgaattact 240
tcctcagaa gagtgtaaag aaaagtcaga gatgctataa tagcagctat ttttaattggc 300
aagtgccact gtggaaagag ttctgtgtg tgctgaagtt ctgaagggca gtcaaattca 360
tcagcatggg ctgtttggtg caaatgcaaa agcacaggtc tttttagcat gctggctctc 420
cccgtgtcct tatgcaaata atcgtcttct tctaaatttc tcttaggctt ctttttccaa 480
agttcttctt ggtttgtgat gtctttcttg ctttccatta attctataaa atagtatggc 540
ttcagccacc cactcttcgc cttagcttga ccgtgagctc cggctgccgc tg 592

<210> 343
<211> 382
<212> DNA
<213> Homo sapien

<400> 343
ttcttgacct cctcctcctt caagctcaaa caccacctcc cttattcagg accggcactt 60
cttaatgttt gtggctttct ctccagcctc tcttaggagg ggtaatgggt gagttggcat 120
cttgtaactc tcctttctcc tttcttcccc tttctctgcc cgcctttccc atcctgctgt 180
agacttcttg attgtcagtc tgtgtcacat ccagtgattg ttttggtttc tgttcccttt 240
ctgactgccc aaggggctca gaaccccagc aatcccttcc tttcactacc ttcttttttg 300
ggggtagttg gaagggactg aaattgtggg gggaaggtag gaggcacatc aataaagagg 360
aaaccaccaa gctgaaaaaa aa 382

<210> 344
<211> 536
<212> DNA
<213> Homo sapien

<400> 344

ctgggcctga	agctgtaggg	taaatcagag	gcaggcttct	gagtgatgag	agtcctgaga	60
caataggcca	cataaacttg	gctggatgga	acctcacaat	aagggtggtca	cctcttgttt	120
gtttaggggg	atgccaaagg	taaggccagc	tcagttatat	gaagagaagc	agaacaaaca	180
agtcctttcag	agaaatggat	gcaatcagag	tgggatcccg	gtcacatcaa	ggtcacactc	240
caccttcatg	tgctggaatg	gttgccagg	cagaaaaatc	cacccttac	gagtgcggct	300
tcgaccctat	atcccccgcc	cgcgtccctt	tctccataaa	attcttctta	gtagctatta	360
ccttcttatt	atttgatcta	gaaattgccc	tccttttacc	cctaccatga	gccctacaaa	420
caactaacct	gccactaata	gttatgtcat	ccctcttatt	aatcatcatc	ctagccctaa	480
gtctggccta	tgagtgacta	caaaaaggat	tagactgagc	cgaataacaa	aaaaaa	536

<210> 345

<211> 251

<212> DNA

<213> Homo sapien

<400> 345

accttttgag	gtctctctca	ccacctccac	agccaccgtc	accgtgggat	gtgctggatg	60
tgaatgaagc	ccccatcttt	gtgcctcctg	aaaagagagt	ggaagtgtcc	gaggactttg	120
gcgtgggcca	ggaaatcaca	tcctacactg	cccaggagcc	agacacattt	atggaacaga	180
aaataacata	tcggatttgg	agagacactg	ccaactggct	ggagattaat	ccggacactg	240
gtgccatttc	c					251

<210> 346

<211> 282

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(282)

<223> n = A,T,C or G

<400> 346

cgcgtctctg	acactgtgat	catgacaggg	gttcaaacag	aaagtgcctg	ggccctcctt	60
ctaagtcttg	ttacccaaaa	aaggaaaaag	aaaagatctt	ctcagttaca	aattctggga	120
agggagacta	tacctggctc	ttgccctaag	tgagaggtct	tcctcccgc	acaaaaaat	180
agaaaggctt	tctatttcac	tggcccagg	agggggaagg	agagtaactt	tgagtctgtg	240
ggtctcattt	ccaaggtgc	cttcaatgct	catnaaaacc	aa		282

<210> 347

<211> 201

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(201)

<223> n = A,T,C or G

aatcgag

908

<210> 351

<211> 472

<212> DNA

<213> Homo sapien

<400> 351

ccagttat	ttt	gcaagtgg	ta	agagcctatt	taccataaat	aataactaaga	accaactcaa	60
gtcaaacctt	aatgccattg	ttattgtgaa	ttaggattaa	gtagtaattt	tcaaaattca			120
cattaacttg	attttaaaat	cagwtttgyg	agtcatttac	cacaagctaa	atgtgtacac			180
tatgataaaa	acaaccattg	tattcctggt	tttctaaaca	gtcctaattt	ctaactgt			240
atatatcctt	cgacatcaat	gaactttgtt	ttcttttact	ccagtaataa	agtaggcaca			300
gatctgtcca	caacaaactt	gccctctcat	gccttgctc	tcaccatgct	ctgctccagg			360
tcagccccct	tttggcctgt	ttgttttgtc	aaaaacctaa	tctgcttctt	gcttttcttg			420
gtaatatata	tttagggaag	atgttgcttt	gcccacacac	gaagcaaagt	aa			472

<210> 352

<211> 251

<212> DNA

<213> Homo sapien

<400> 352

ctcaaagcta	atctctcggg	aatcaaacca	gaaaagggca	aggatcttag	gcatgggtgga	60
tgtggataag	gccaggtcaa	tggctgcaag	catgcagaga	aagaggtaca	tcggagcgtg	120
caggctgcgt	tccgtcctta	cgatgaagac	cacgatgcag	tttccaaaca	ttgccactac	180
atacatggaa	aggaggggga	agccaaccca	gaaatgggct	ttctctaate	ctgggataacc	240
aataagcaca	a					251

<210> 353

<211> 436

<212> DNA

<213> Homo sapien

<400> 353

tttttttttt	tttttttttt	tttttttaca	caatgcagtc	atatttttat	tgagtatgtg	60
cacattatgg	tattattact	atactgatta	tatttatcat	gtgacttcta	attaraaaat	120
gtatccaaaa	gcaaaacagc	agatatataa	aattaaagag	acagaagata	gacattaaca	180
gataaggcaa	cttatacatt	gacaatccaa	atccaatata	tttaaacatt	tgggaaatga	240
gggggacaaa	tgggaagccar	atcaaatttg	tgtaaaacta	ttcagtatgt	ttcccttgct	300
tcatgtctga	raaggctctc	ccttcaatgg	ggatgacaaa	ctccaaatgc	cacacaaatg	360
ttaacagaat	actagattca	cactggaacg	ggggtaaaaga	agaaattatt	ttctataaaa	420
gggctcctaa	tgtagt					436

<210> 354

<211> 854

<212> DNA

<213> Homo sapien

<400> 354

ccttttctag	ttcaccagtt	ttctgcaagg	atgctgggtta	gggagtgtct	gcaggaggag	60
------------	------------	------------	-------------	------------	------------	----

caagtctgaa	accaaatacta	ggaaacatag	gaaacgagcc	aggcacaggg	ctgggtgggccc	120
atcagggacc	acccttttggg	ttgatatttt	gcttaaatctg	catcttttga	gtaagatcat	180
ctggcagtag	aagctgttct	ccaggtacat	ttctctagct	catgtacaaa	aacatcctga	240
aggactttgt	caggtgcctt	gctaaaagcc	agatgcgttc	ggcacttcct	tggctctgagg	300
ttaattgcac	acctacaggc	actgggtcca	tgctttcaag	tattttgtcc	tcacttttagg	360
gtgagtga	gatccccatt	ataggagcac	ttgggagaga	tcataataaa	gctgactcct	420
gagtacatgc	agtaatgggg	tagatgtgtg	tgggtgtgtct	tcattcctgc	aagggtgcctt	480
gttagggagt	gtttccagga	ggaacaagtc	tgaaaccaat	catgaaataa	atggtaggtg	540
tgaactggaa	aactaattca	aaagagagat	cgtgatataca	gtgtgggtga	tacaccttgg	600
caatatggaa	ggctctaatt	tgcccatatt	tgaaataata	attcagcttt	ttgtaataca	660
aaataacaaa	ggattgagaa	tcattggtgtc	taatgtataa	aagaccaggg	aaacataaat	720
atatcaactg	cataaatgta	aaatgcatgt	gacccaagaa	ggcccaaaag	tggcagacaa	780
cattgtaccc	attttccctt	ccaaaatgtg	agcggcgggc	ctgctgcttt	caaggctgtc	840
acacgggatg	tcag					854

<210> 355

<211> 676

<212> DNA

<213> Homo sapien

<400> 355

gaaattaagt	atgagctaaa	ttccctgtta	aaacctctag	gggtgacaga	tctcttcaac	60
caggtcaaag	ctgatctttc	tggatgtca	ccaaccaagg	gcctatatatt	atcaaaagcc	120
atccacaagt	catacctgga	tgtcagcgaa	gagggcacgg	aggcagcagc	agccactggg	180
gacagcatcg	ctgtaaaaaag	cctaccaatg	agagctcagt	tcaaggcgaa	ccacccttcc	240
ctgttcttta	taaggcacac	tcataccaac	acgatcctat	tctgtggcaa	gcttgccctct	300
ccctaatacag	atgggggttga	gtaaggctca	gagttgcaga	tgaggtgcag	agacaatcct	360
gtgactttcc	cacggccaaa	aagctgttca	cacctcacgc	acctctgtgc	ctcagtttgc	420
tcactctgcaa	aatagggtcta	ggatttcttc	caaccatttc	atgagttgtg	aagctaaggc	480
tttgttaatc	atggaaaaag	gtagacttat	gcagaaagcc	tttctggctt	tcttatctgt	540
ggtgtctcat	ttgagtgtcg	tccagtgaca	tgatcaagtc	aatgagtaaa	attttaaggg	600
attagatttt	cttgacttgt	atgtatctgt	gagatcttga	ataagtgacc	tgacatctct	660
gcttaaaagaa	aaccag					676

<210> 356

<211> 574

<212> DNA

<213> Homo sapien

<400> 356

tttttttttt	tttttcagga	aaacattctc	ttacttttatt	tgcattctcag	caaaggttct	60
catgtggcac	ctgactggca	tcaaaccaaa	gttcgttaggc	caacaaagat	gggccactca	120
caagcttccc	atttgtagat	ctcagtgcct	atgagtatct	gacacctgtt	cctctcttca	180
gtctcttagg	gaggcttaaa	tctgtctcag	gtgtgctaag	agtgccagcc	caaggkgttc	240
aaaagtccac	aaaactgcag	tctttgctgg	gatagtaagc	caagcagtgc	ctggacagca	300
gagttctttt	cttgggcaac	agataaccag	acaggactct	aatcgtgtct	ttattcaaca	360
ttcttctgtc	tctgcctaga	ctggaataaa	aagccaatct	ctctcgtggc	acaggggaag	420
agatacaagc	togtttacat	gtgatagatc	taacaaaggc	atctaccgaa	gtctggtctg	480
gatagacggc	acagggagct	cttaggtcag	cgctgctggg	tggaggacat	tcctgagtcc	540
agctttgcag	cctttgtgca	acagtacttt	ccca			574

<210> 357
 <211> 393
 <212> DNA
 <213> Homo sapien

<400> 357
 tttttttttt tttttttttt tttttttttt tacagaatat aratgcttta tcaactgkact 60
 taatatggkg kcttggtcac tatacttaaa aatgcaccac tcataaatat ttaattcagc 120
 aagccacaac caaracttga ttttatcaac aaaaaccctt aaatataaac ggsaaaaaag 180
 atagatataa ttattccagt ttttttaaaa cttaaaarat attccattgc cgaattaara 240
 araarataag tggtatatgg aaagaagggc attcaagcac actaaaraaa cctgaggkaa 300
 gcataatctg tacaaaatta aactgtcctt tttggcattt taacaaattt gcaacgktct 360
 tttttttctt tttctgtttt tttttttttt tac 393

<210> 358
 <211> 630
 <212> DNA
 <213> Homo sapien

<400> 358
 acagggtaaa caggaggatc cttgctctca cggagcttac attctagcag gaggacaata 60
 ttaatgttta taggaaaatg atgagtttat gacaaaggaa gtagatagtg ttttacaaga 120
 gcatagagta gggaagctaa tccagcacag ggaggtcaca gagacatccc taaggaagtg 180
 gagtttaaac tgagagaagc aagtgcctaa actgaaggat gtgttgaaga agaagggaga 240
 gtagaacaat ttgggcagag ggaaccttat agaccctaag gtgggaaggt tcaaagaact 300
 gaaagagagc tagaacagct ggagccgttc tccggtgtaa agaggagtca aagagataag 360
 attaaagatg tgaagattaa gatcttggtg gcattcaggg attggcactt ctacaagaaa 420
 tcaactgaagg gagtaatgtg acattacttt tcaacttcagg atggccattc taactccagg 480
 gggtagactg gactaggtaa gactggaggc aggtagacct cttctaaggc ctgcatagtg 540
 gaaagacaaa aataagtggg gaaattcagg ggatagtgaa aatcagtagg acttaatgag 600
 caagccagag gttcctccac aacaaccagt 630

<210> 359
 <211> 620
 <212> DNA
 <213> Homo sapien

<400> 359
 acagcattcc aaaatataca tctagagact aarrgtaaat gctctatagt gaagaagtaa 60
 taattaaaaa atgctactaa tatagaaaat ttataatcag aaaaataaat attcagggag 120
 ctcaccagaa gaataaagtg ctctgccagt tattaaagga ttactgctgg tgaattaaat 180
 atggcattcc ccaagggaaa tagagagatt cttctggatt atgttcaata tttatttcac 240
 aggattaact gttttaggaa cagatataaa gcttcgccac ggaagagatg gacaaagcac 300
 aaagacaaca tgatacctta ggaagcaaca ctaccctttc aggcataaaa tttggagaaa 360
 tgcaacatta tgcttcatga ataatatgta gaaagaaggt ctgatgaaaa tgacatcctt 420
 aatgtaagat aactttataa gaattctggg tcaaataaaa ttctttgaag aaaacatcca 480
 aatgtcattg acttatcaaa tactatcttg gcatataacc tatgaaggca aaactaaaca 540
 aacaaaaagc tcacaccaa caaaaccatc aacttatttt gtattctata acatacgaga 600
 ctgtaaagat gtgacagtgt 620

<210> 360

<211> 431
 <212> DNA
 <213> Homo sapien

<400> 360
 aaaaaaaaaa agccagaaca acatgtgata gataaatatga ttggctgcac acttccagac 60
 tgatgaatga tgaacgtgat ggactattgt atggagcaca tcttcagcaa gagggggaaa 120
 tactcatcat ttttggccag cagttgtttg atcaccaaac atcatgccag aatactcagc 180
 aaaccttctt agctcttgag aagtcaaagt cggggggaat ttattcctgg caattttaat 240
 tggactcctt atgtgagagc agcggctacc cagctggggt ggtggagcga acccgctcact 300
 agtggacatg cagtggcaga gtccttggtg accacctaga ggaatacaca ggcacatgtg 360
 tgatgccaag cgtgacacct gtgacactca aatttgtctt gtttttgtct ttcgggtgtgt 420
 agattcttag t 431

<210> 361
 <211> 351
 <212> DNA
 <213> Homo sapien

<400> 361
 aacttgattt ccgatcaaaa gaatcatcat ctttaccttg acttttcagg gaattactga 60
 actttcttct cagaagatag ggcacagcca ttgccttggc ctcacttgaa gggctctgcat 120
 ttgggtcctc tggctctctg ccaagtttcc cagccactcg agggagaaat atcgggaggt 180
 ttgacttctt ccggggcttt cccgagggct tcaccgtgag ccctgcggcc ctcagggctg 240
 caatcctgga ttcaatgtct gaaacctcgc tctctgcctg ctggacttct gaggccgtca 300
 ctgccactct gtcctccagc tctgacagct cctcatctgt ggtcctgttg t 351

<210> 362
 <211> 463
 <212> DNA
 <213> Homo sapien

<400> 362
 acttcatcag gccataatgg gtgcctcccg tgagaatcca agcacctttg gactgcgcga 60
 tgtagatgag ccggctgaag atcttgcgca tgcgcggctt cagggcgaag ttcttggcgc 120
 ccccggtcac agaaatgacc aggttgggtg ttttcagggt ccagtgtctg gtcagcagct 180
 cgtaaaggat ttccgcgtcc gtgtcgcagg acagacgtat atacttccct ttcttcccca 240
 gtgtctcaaa ctgaatatcc ccaaaggcgt cggtaggaaa ttccttgggt tgtttcttgt 300
 agttccattt ctcacttttg ttgatctggg tgccttccat gtgctggctc tgggcatagc 360
 cacacttgca cacattctcc ctgataagca cgatggtgtg gacaggaagg aaggatttca 420
 ttgagcctgc ttatggaaac tggatttgtt agcttaata gac 463

<210> 363
 <211> 653
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(653)
 <223> n = A,T,C or G

<400> 363

acccccgagt	ncctgnctgg	catactgnga	acgaccaacg	acacacccaa	gctcggcctc	60
ctcttgngga	ttctgggtga	catcttcatg	aatggcaacc	gtgccagwga	ggctgtcctc	120
tgggaggcac	tacgcaagat	gggactgcgt	cctgggggtga	gacatcctct	ccttggagat	180
ctaacgaaac	ttctcaccta	tgagttgtaa	agcagaaata	cctgnactac	agacgagtgc	240
ccaacagcaa	ccccccggaa	gtatgagttc	ctctrgggcc	tccgttcccta	ccatgagasc	300
tagcaagatg	naagtgttga	gantcattgc	agaggttcag	aaaagagacc	cntcgtgact	360
ggtctgcaca	gttcatggag	gctgcagatg	aggccttggga	tgctctggat	gctgctgcag	420
ctgaggccga	agccccggct	gaagcaagaa	cccgcattggg	aattggagat	gaggctgtgt	480
ntgggccctg	gagctgggat	gacattgagt	ttgagctgct	gacctgggat	gaggaaggag	540
attttgagga	tccttgggtcc	agaattccat	ttaccttctg	ggccagatac	caccagaatg	600
cccgtccag	attccctcag	acctttgccg	gtcccattat	tggtcstggt	ggt	653

<210> 364

<211> 401

<212> DNA

<213> Homo sapien

<400> 364

actagaggaa	agacgttaaa	ccactctact	accacttgtg	gaactctcaa	agggtaaatg	60
acaaagccaa	tgaatgactc	taaaaacaat	atttacattt	aarggtttgt	agacaataaa	120
aaaacaaggt	ggatagatct	agaattgtaa	cattttaaga	aaaccatagc	atttgacaga	180
tgagaaagct	caattataga	tgcaaagtta	taactaaact	actatagtag	taaagaaata	240
catttcacac	ccttcatata	aatttcaact	cttggcttga	ggcactccat	aaaatgtatc	300
acgtgcatag	taaatcttta	tatttgctat	ggcgttgcac	tagaggactt	ggactgcaac	360
aagtgyatgc	gcggaaaaatg	aatcttctct	caatagccca	g		401

<210> 365

<211> 356

<212> DNA

<213> Homo sapien

<400> 365

ccagtgtcat	atttgggctt	aaaatttcaa	gaagggcact	tcaaattggct	ttgcatttgc	60
atgtttcagt	gctagagcgt	aggaatagac	cctggcgctc	actgtgagat	gttcttcagc	120
taccagagca	tcaagtctct	gcagcaggtc	attcttgggt	aaagaaatga	cttcacaaa	180
ctctccatcc	cctggccttg	gcttoggcct	tgcgttttcg	gcacatcttc	cgttaatggt	240
gactgtcacg	atgtgtatag	tacagtttga	caagcctggg	tccatacaga	ccgctggaga	300
acattcggca	atgtcccctt	tgtagccagt	ttcttcttcg	agctcccgga	gagcag	356

<210> 366

<211> 1851

<212> DNA

<213> Homo sapien

<400> 366

tcatacccat	tgccagcagc	ggcaccgtta	gtcaggtttt	ctgggaatcc	cacatgagta	60
cttccgtgtt	cttcattctt	cttcaatagc	cataaatctt	ctagctctgg	ctggctgttt	120
tcacttcctt	taagcctttg	tgactcttcc	tctgatgtca	gctttaagtc	ttgttctgga	180
ttgctgtttt	cagaagagat	ttttaacatc	tgtttttctt	tgtagtcaga	aagtaactgg	240

<400> 368

gggtcgccca	ggggsgcgt	gggctttcct	cgggtgggtg	tgggttttcc	ctgggtgggg	60
tgggctgggc	trgaatcccc	tgctgggggt	ggcagggtttt	ggctgggatt	gacttttytc	120
ttcaaacaga	ttggaaaccc	ggagttacct	gctagttggg	gaaactgggt	ggtagacgcg	180
atctgttggc	tactactggc	ttctcctggc	tgtaaaaagc	agatgggtgg	tgaggttgat	240
tccatgccgg	ctgcttcttc	tgtgaagaag	ccatttggtc	tcaggagcaa	gatgggcaag	300
tggtgctgcc	gttgcttccc	ctgctgcagg	gagagcggca	agagcaacgt	gggcacttct	360
ggagaccacg	acgactctgc	tatgaagaca	ctcaggagca	agatgggcaa	gtggtgccgc	420
cactgcttcc	cctgctgcag	ggggagtggc	aagagcaacg	tgggcgcttc	tggagaccac	480
gacgaytctg	ctatgaagac	actcaggaac	aagatgggca	agtgggtgctg	ccactgcttc	540
ccctgctgca	gggggagcrg	caagagcaag	gtgggcgctt	ggggagacta	cgatgacagt	600
gccttcatgg	agcccaggta	ccacgtccgt	ggagaagatc	tggacaagct	ccacagagct	660
gcctgggtggg	gtaaagtccc	cagaaaggat	ctcatcgcca	tgctcaggga	cactgacgtg	720
aacaagaagg	acaagcaaaa	gaggactgct	ctacatctgg	cctctgcca	tgggaattca	780
gaagtagtaa	aactcstgct	ggacagacga	tgtcaactta	atgtccttga	caacaaaaag	840
aggacagctc	tgayaaaggc	cgtacaatgc	caggaagatg	aatgtgcggt	aatgttgctg	900
gaacatggca	ctgatccaaa	tattccagat	gagtatggaa	ataccactct	rcactaygct	960
rtctayaatg	aagataaatt	aatggccaaa	gcactgctct	tatayggtgc	tgatatcgaa	1020
tcaaaaaaca	aggtatagat	ctactaattt	tatcttcaaa	atactgaaat	gcattcattt	1080
taacattgac	gtgtgtaagg	gccagtcttc	cgtatttgga	agctcaagca	taacttgaat	1140
gaaaatattt	tgaaatgacc	taattatctm	agactttatt	ttaaatattg	ttattttcaa	1200
agaagcatta	gaggttacag	tttttttttt	ttaaatgcac	ttctggtaaa	tacttttggt	1260
gaaaacactg	aatttgtaaa	aggtaatact	tactattttt	caatttttcc	ctcctaggat	1320
ttttttcccc	taatgaatgt	aagatggcaa	aatttgccct	gaaatagggt	ttacatgaaa	1380
actccaagaa	aagttaaaca	tgtttcagtg	aatagagatc	ctgctccttt	ggcaagttcc	1440
taaaaaacag	taatagatac	gaggtgatgc	gcctgtcagt	ggcaagggtt	aagatatattc	1500
tgatctcgtg	cc					1512

<210> 369

<211> 1853

<212> DNA

<213> Homo sapien

<400> 369

gggtcgccca	ggggsgcgt	gggctttcct	cgggtgggtg	tgggttttcc	ctgggtgggg	60
tgggctgggc	trgaatcccc	tgctgggggt	ggcagggtttt	ggctgggatt	gacttttytc	120
ttcaaacaga	ttggaaaccc	ggagttacct	gctagttggg	gaaactgggt	ggtagacgcg	180
atctgttggc	tactactggc	ttctcctggc	tgtaaaaagc	agatgggtgg	tgaggttgat	240
tccatgccgg	ctgcttcttc	tgtgaagaag	ccatttggtc	tcaggagcaa	gatgggcaag	300
tggtgctgcc	gttgcttccc	ctgctgcagg	gagagcggca	agagcaacgt	gggcacttct	360
ggagaccacg	acgactctgc	tatgaagaca	ctcaggagca	agatgggcaa	gtggtgccgc	420
cactgcttcc	cctgctgcag	ggggagtggc	aagagcaacg	tgggcgcttc	tggagaccac	480
gacgaytctg	ctatgaagac	actcaggaac	aagatgggca	agtgggtgctg	ccactgcttc	540
ccctgctgca	gggggagcrg	caagagcaag	gtgggcgctt	ggggagacta	cgatgacagy	600
gccttcatgg	akcccaggta	ccacgtccrt	ggagaagatc	tggacaagct	ccacagagct	660
gcctgggtggg	gtaaagtccc	cagaaaggat	ctcatcgcca	tgctcaggga	cackgaygtg	720
aacaagargg	acaagcaaaa	gaggactgct	ctacatctgg	cctctgcca	tgggaattca	780
gaagtagtaa	aactcstgct	ggacagacga	tgtcaactta	atgtccttga	caacaaaaag	840
aggacagctc	tgayaaaggc	cgtacaatgc	caggaagatg	aatgtgcggt	aatgttgctg	900
gaacatggca	ctgatccaaa	tattccagat	gagtatggaa	ataccactct	rcactaygct	960

rtctayaatg	aagataaatt	aatggccaaa	gcactgctct	tataygggtgc	tgatatcgaa	1020
tcaaaaaaca	agcatggcct	cacaccactg	ytacttggtr	tacatgagca	aaaacagcaa	1080
gtsgtgaaat	ttttaatyaa	gaaaaaagcg	aattttaaatt	gcrctggata	gatatggaag	1140
ractgctctc	atacttgctg	tatgttgtgg	atcagcaagt	atagtcagcc	ytctacttga	1200
gcaaaaatrtt	gatgtatctt	ctcaagatct	ggaaagacgg	ccagagagta	tgctgtttct	1260
agtcacatc	atgtaatttg	ccagttactt	tctgactaca	aagaaaaaca	gatgttaaaa	1320
atctcttctg	aaaacagcaa	tccagaacaa	gacttaaagc	tgacatcaga	ggaagagtca	1380
caaaggctta	aaggaagtga	aaacagccag	ccagaggcat	ggaaactttt	aaattttaaac	1440
ttttggttta	atgttttttt	tttttgctt	aataatatta	gatagtccca	aatgaaatwa	1500
cctatgagac	taggctttga	gaatcaatag	attctttttt	taagaatctt	ttggctagga	1560
gcggtgtctc	acgcctgtaa	ttccagcacc	ttgagaggct	gaggtgggca	gatcacgaga	1620
tcaggagatc	gagaccatcc	tggctaacac	ggtgaaaccc	catctctact	aaaaatacaa	1680
aaacttagct	gggtgtgggtg	gcggtgcct	gtagtcccag	ctactcagga	rgctgaggca	1740
ggagaatggc	atgaacccgg	gaggtggagg	ttgcagttag	ccgagatccg	ccactacact	1800
ccagcctggg	tgacagagca	agactctgtc	tcaaaaaaaaa	aaaaaaaaaaa	aaa	1853

<210> 370

<211> 2184

<212> DNA

<213> Homo sapien

<400> 370

ggcacgagaa	ttaaaaccct	cagcaaaaaca	ggcatagaag	ggacatacct	taaagtaata	60
aaaaccacct	atgacaagcc	cacagccaac	ataatactaa	atgggggaaaa	gttagaagca	120
tttcctctga	gaactgcaac	aataaaatata	aggatgctgg	attttgtcaa	atgccttttc	180
tgtgtctgtt	gagatgctta	tgtgactttg	cttttaattc	tgtttatgtg	attatcacat	240
ttattgactt	gcctgtgtta	gaccggaaga	gctgggggtgt	ttctcaggag	ccaccgtgtg	300
ctgcggcagc	ttcgggataa	cttgaggctg	catcactggg	gaagaaacac	aytcctgtcc	360
gtggcgctga	tggctgagga	cagagcttca	gtgtggcttc	tctgcgactg	gcttcttcgg	420
ggagtctctc	cttcatagtt	catccatatg	gctccagagg	aaaattatat	tattttgtta	480
tggatgaaga	gtattacgtt	gtgcagatat	actgcagtgt	cttcatctct	tgatgtgtga	540
ttgggtaggt	tccaccatgt	tgcgcagat	gacatgattt	cagtacctgt	gtctggctga	600
aaagtgtttg	tttgtgaatg	gatattgtgg	ttcttggtatc	tcattctctg	tgggtggaca	660
gctttctcca	ccttgcctgga	agtgcacctg	tgtccagaag	tttgatggct	gaggagtata	720
ccatcgtgca	tgcactcttc	atttctctgca	tttcttctct	cctggatgga	cagggggagc	780
ggcaagagca	acgtgggcac	ttctggagac	cacaacgact	cctctgtgaa	gacgcttggg	840
agcaagaggt	gcaagtgggtg	ctgccactgc	ttcccctgct	gcagggggagc	ggcaagagca	900
acgtggtcgc	ttggggagac	tacgatgaca	gcgccttcat	ggatcccagg	taccacgtcc	960
atggagaaga	tctggacaag	ctccacagag	ctgcctgggtg	gggtaaagtc	cccagaaagg	1020
atctcatcgt	catgctcagg	gacacggatg	tgaacaagag	ggacaagcaa	aagaggactg	1080
ctctacatct	ggcctctgcc	aatgggaatt	cagaagtagt	aaaactcgtg	ctggacagac	1140
gatgtcaact	taatgtcctt	gacaacaaaa	agaggacagc	tctgacaaag	gccgtacaat	1200
gccaggaaga	tgaatgtgcg	ttaatgttgc	tggaaacatgg	cactgatcca	aatattccag	1260
atgagtatgg	aaataccact	ctacactatg	ctgtctacaa	tgaagataaa	ttaatggcca	1320
aagcactgct	cttatacggg	gctgatatcg	aatcaaaaaa	caagcatggc	ctcacaccac	1380
tgctacttgg	tatacatgag	caaaaacagc	aagtggtgaa	atttttaatc	aagaaaaaag	1440
cgaattttaa	tgcgctggat	agatatggaa	gaactgctct	catacttgct	gtatgttgtg	1500
gatcagcaag	tatagtcagc	cctctacttg	agcaaaatgt	tgatgtatct	tctcaagatc	1560
tggaaagacg	gccagagagt	atgctgtttc	tagtcatcat	catgtaattt	gccagttact	1620
ttctgactac	aaagaaaaac	agatgttaaa	aatctcttct	gaaaacagca	atccagaaca	1680
agacttaaa	ctgacatcag	aggaagagtc	acaaaggctt	aaaggaagtg	aaaacagcca	1740

```

gccagaggca tggaaacttt taaattttaa ctttttggtt aatgtttttt ttttttgctt 1800
taataatatt agatagtcctt aaatgaaatw acctatgaga ctaggctttg agaatacaata 1860
gattcttttt ttaagaatct tttggctagg agcgggtgtct caccgctgta attccagcac 1920
cttgagaggc tgaggtgggc agatcacgag atcaggagat cgagaccatc ctgggctaaca 1980
cggtgaaacc ccatctctac taaaaataca aaaacttagc tgggtgtggt ggcggttgcc 2040
tgtagtccca gctactcagg argctgaggc aggagaatgg catgaaccgc ggaggtggag 2100
gttgagtgta gccgagatcc gccactacac tccagcctgg gtgacagagc aagactctgt 2160
ctcaaaaaaa aaaaaaaaaa aaaa 2184

```

<210> 371

<211> 1855

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(1855)

<223> n = A,T,C or G

<400> 371

```

tgcacgcac ggccagtgtc tgtgccacgt aactgaagc cccctgagat gtgcacgccg 60
cacgcgcac ttgcacgcgc ggcagcggct tggctggctt gtaacggctt gcacgcgcac 120
gccgcccccg cataaccgtc agactggcct gtaacggctt gcagggcgac gccgcacgcg 180
cgtaacggct tggctgccct gtaacggctt gcacgtgcat gctgcacgug cgttaacggc 240
ttggctggca ttagccgctt tggcttggct ttgcatttct tgetkggctk ggcggttgkty 300
tcttggattg acgttctctc cttggatkgc cgttctctcc ttggatkgac gtttctytyt 360
tcgcgttctt ttgctggact tgacctttty tctgctgggt ttggcattcc tttgggggtg 420
gctgggtgtt ttctccgggg gggkktgccc tctctggggg gggcggtgggk cgcgcccgag 480
gggcgtgggc tttccccggg tgggtgtggg ttttctggg gtgggggtggg ctgtgctggg 540
atccccctgc tggggttggc agggattgac ttttttcttc aaacagattg gaaaccggga 600
gtaacntgct agttggtgaa actggttggg agacgcgac tgctggtact actgtttctc 660
ctggctgtta aaagcagatg gtggctgagg ttgattcaat gccggctgct tcttctgtga 720
agaagccatt tggctctcag agcaagatgg gcaagtgggt cgcactgct tcccctgctg 780
caggggggag ggcaagagca acgtgggcac ttctggagac cacaacgact cctctgtgaa 840
gacgcttggg agcaagaggt gcaagtgggt ctgcccactg cttccccctg tgcaaggagg 900
cggcaagagc aacgtggkcg cttggggaga ctacgatgac agcgccttca tggakcccag 960
gtaccacgtc crtggagaag atctggacaa gctccacaga gctgcctggt ggggtaaagt 1020
ccccagaaag gatctcatcg tcatgctcag ggacactgay gtgaacaaga rggacaagca 1080
aaagaggact gctctacatc tggcctctgc caatgggaat tcagaagtag taaaactcgt 1140
gctggacaga cgatgtcaac ttaatgtcct tgacaacaaa aagaggacag ctctgacaaa 1200
ggcgtacaa tgccaggaag atgaatgtgc gttaatgttg ctggaacatg gcaatgatcc 1260
aaatattcca gatgagtatg gaaataccac tctacactat gctgtctaca atgaagataa 1320
attaatggcc aaagcactgc tcttatacgg tgetgatata gaatcaaaaa acaaggtata 1380
gatctactaa ttttatcttc aaaatactga aatgcattca ttttaacatt gacgtgtgta 1440
agggccagtc ttccgtatctt ggaagctcaa gcataacttg aatgaaaata ttttgaaatg 1500
acctaattat ctaagacttt attttaaaata ttgttatttt caaagaagca ttagagggtg 1560
cagttttttt tttttaaatg cacttctggt aaatactttt gttgaaaaca ctgaatttgt 1620
aaaaggtaat acttactatt tttcaatttt tccctcctag gatttttttc ccctaataaa 1680
tgtaagatgg caaaatttgc cctgaaatag gttttacatg aaaactccaa gaaaagttaa 1740
acatgtttca gtgaatagag atcctgctcc tttggcaagt tcctaaaaaa cagtaataga 1800
tacgaggtga tgcgcctgtc agtggcaagg ttttaagatat ttctgatctc gtgcc 1855

```


<210> 372
 <211> 1059
 <212> DNA
 <213> Homo sapien

<400> 372

gcaacgtggg	cacttctgga	gaccacaacg	actcctctgt	gaagacgctt	gggagcaaga	60
ggtgcaagt	gtgctgccc	ctgcttcccc	tgctgcagg	gagcggcaag	agcaacgtgg	120
gcgcttgrg	agactmcgat	gacagygcct	tcatggagcc	caggtaccac	gtccgtggag	180
aagatctgga	caagctccac	agagctgccc	tgggtgggta	aagtccccag	aaaggatctc	240
atcgatcat	tcagggacac	tgaygtgaac	aagarggaca	agcaaaagag	gactgctcta	300
catctggcct	ctgccaatgg	gaattcagaa	gtagtataac	tcstgctgga	cagacgatgt	360
caacttaat	tccttgacaa	caaaaagagg	acagctctga	yaaaggccgt	acaatgccag	420
gaagatgaat	gtgcgttaat	gttgctggaa	catggcactg	atccaaatat	tccagatgag	480
tatggaaata	ccactctrca	ctaygctrct	tayaatgaag	ataaattaat	ggccaaagca	540
ctgctcttat	ayggtgctga	tatcgaatca	aaaaacaagg	tatagatcta	ctaattttat	600
cttcaaaata	ctgaaatgca	ttcattttta	cattgacgtg	tgttaagggcc	agtcttccgt	660
atttggaa	tcaagcataa	cttgaatgaa	aatatattga	aatgacctaa	ttatctaaga	720
ctttatttta	aatattgtta	ttttcaaaga	agcattagag	ggtacagttt	ttttttttta	780
aatgcacttc	tggtaaatac	ttttgttgaa	aacactgaat	ttgtaaaagg	taatacttac	840
tatttttcaa	tttttccctc	ctaggatttt	tttcccttaa	tgaatgtaag	atggcaaaat	900
ttgccctgaa	ataggtttta	catgaaaact	ccaagaaaag	ttaaacatgt	ttcagtgaat	960
agagatcctg	ctcctttggc	aagttcctaa	aaaacagtaa	tagatacgag	gtgatgcgcc	1020
tgtcagtggc	aaggtttaag	atatttctga	tctcgtgcc			1059

<210> 373
 <211> 1155
 <212> DNA
 <213> Homo sapien

<400> 373

atggtggttg	aggttgattc	catgccggct	gcctcttctg	tgaagaagcc	atttggctct	60
aggagcaaga	tgggcaagt	gtgctgccgt	tgttccccct	gctgcaggga	gagcggcaag	120
agcaacgtgg	gcacttctgg	agaccacgac	gactctgcta	tgaagacact	caggagcaag	180
atgggcaagt	ggtgccgcca	ctgcttcccc	tgctgcagg	ggagtggcaa	gagcaacgtg	240
ggcgcttctg	gagaccacga	cgactctgct	atgaagacac	tcaggaacaa	gatgggcaag	300
tgggtgctgcc	actgcttccc	ctgctgcagg	gggagcggca	agagcaagg	gggcgcttgg	360
ggagactacg	atgacagtgc	cttcatggag	cccagggtacc	acgtccgtgg	agaagatctg	420
gacaagctcc	acagagctgc	ctgggtgggt	aaagtcccca	gaaaggatct	catcgatcat	480
ctcaggga	ctgacgtgaa	caagaaggac	aagcaaaaaga	ggactgctct	acatctggcc	540
tctgccaatg	ggaattcaga	agtagtaaaa	ctcctgctgg	acagacgatg	tcaacttaat	600
gtccttgaca	acaaaaagag	gacagctctg	ataaaggccg	tacaatgcca	ggaagatgaa	660
tgtgcgttaa	tgttgctgga	acatggcact	gatccaaata	ttccagatga	gtatggaaat	720
accactctgc	actacgctat	ctataatgaa	gataaattaa	tgcccaaagc	actgctctta	780
tatggtgctg	atatcgaatc	aaaaaacaag	catggcctca	caccactgtt	acttgggtga	840
catgagcaaa	aacagcaagt	cgtgaaat	ttaatcaaga	aaaaagcgaa	tttaaatgca	900
ctggatagat	atggaaggac	tgtctcata	cttgctgtat	gttggtggatc	agcaagtata	960
gtcagccttc	tacttgagca	aaatattgat	gtatcttctc	aagatctatc	tggaagagag	1020
gccagagagt	atgctgtttc	tagtcacat	catgtaattt	gccagttact	ttctgactac	1080
aaagaaaaac	agatgctaaa	aatctcttct	gaaaacagca	atccagaaaa	tgtctcaaga	1140

accagaaata aataa

1155

<210> 374

<211> 2000

<212> DNA

<213> Homo sapien

<400> 374

atggtggttg	aggttgattc	catgccggct	gcctcttctg	tgaagaagcc	atttggctctc	60
aggagcaaga	tgggcaagt	gtgctgccgt	tgcttcccct	gctgcaggga	gagcggcaag	120
agcaacgtgg	gcacttctgg	agaccacgac	gactctgcta	tgaagacact	caggagcaag	180
atgggcaagt	ggtgccgcca	ctgcttcccc	tgctgcagg	ggagtggcaa	gagcaacgtg	240
ggcgcttctg	gagaccacga	cgactctgct	atgaagacac	tcaggaacaa	gatgggcaag	300
tggtgctgcc	actgcttccc	ctgctgcagg	gggagcgcca	agagcaaggt	gggcgcttgg	360
ggagactacg	atgacagtgc	cttcatggag	cccagggtacc	acgtccgtgg	agaagatctg	420
gacaagctcc	acagagctgc	ctgggtgggt	aaagtcccca	gaaaggatct	catcgtcatg	480
ctcagggaca	ctgacgtgaa	caagaaggac	aagcaaaaga	ggactgctct	acatctggcc	540
tctgccaatg	ggaattcaga	agtagtaaaa	ctcctgctgg	acagacgatg	tcaacttaat	600
gtccttgaca	acaaaaagag	gacagctctg	ataaaggccg	tacaatgcc	ggaagatgaa	660
tgtgcgttaa	tggtgctgga	acatggcact	gatccaaata	ttccagatga	gtatggaaat	720
accactctgc	actacgctat	ctataatgaa	gataaattaa	tggccaaagc	actgctctta	780
tatggtgctg	atatcgaatc	aaaaaacaag	catggcctca	caccactgtt	acttgggtgta	840
catgagcaaa	aacagcaagt	cgtgaaat	ttaatcaaga	aaaaagcgaa	tttaaatgca	900
ctggatagat	atggaaggac	tgctctcata	cttgctgtat	gttggtggatc	agcaagtata	960
gtcagccttc	tacttgagca	aaatat	gtatcttctc	aagatctatc	tggacagacg	1020
gccagagagt	atgctgtttc	tagtcatcat	catgtaattt	gccagttact	ttctgactac	1080
aaagaaaaac	agatgctaaa	aatctcttct	gaaaacagca	atccagaaca	agacttaaag	1140
ctgacatcag	aggaagagtc	acaaagggtc	aaaggcagtg	aaaatagcca	gccagagaaa	1200
atgtctcaag	aaccagaaat	aaataaggat	ggtgatagag	aggttgaaga	agaaatgaag	1260
aagcatgaaa	gtaataatgt	gggattacta	gaaaacctga	ctaattggtgt	cactgctggc	1320
aatggtgata	atggattaat	tctctaaagg	aagagcagaa	cacctgaaaa	tcagcaattt	1380
cctgacaacg	aaagtgaaga	gtatcacaga	atttgcgaat	tagtttctga	ctacaaagaa	1440
aaacagatgc	caaaatactc	ttctgaaaac	agcaaccag	aacaagactt	aaagctgaca	1500
tcagaggaag	agtcacaaag	gcttgagggc	agtgaaaatg	gccagccaga	gctagaaaat	1560
tttatggcta	tcgaagaaat	gaagaagcac	ggaagtactc	atgtcggatt	cccagaaaac	1620
ctgactaatg	gtgccactgc	tggcaatggt	gatgatggat	taattcctcc	aaggaagagc	1680
agaacacctg	aaagccagca	atttcctgac	actgagaatg	aagagtatca	cagtgcagaa	1740
caaaatgata	ctcagaagca	attttgtgaa	gaacagaaca	ctggaatatt	acacgatgag	1800
attctgattc	atgaagaaaa	gcagatagaa	gtgggtgaaa	aaatgaattc	tgagctttct	1860
cttagttgta	agaaagaaaa	agacatcttg	catgaaaata	gtacgttgcg	ggaagaaatt	1920
gccatgctaa	gactggagct	agacacaatg	aaacatcaga	gccagctaaa	aaaaaaaaaa	1980
aaaaaaaaaa	aaaaaaaaaa					2000

<210> 375

<211> 2040

<212> DNA

<213> Homo sapien

<400> 375

atggtggttg	aggttgattc	catgccggct	gcctcttctg	tgaagaagcc	atttggctctc	60
aggagcaaga	tgggcaagt	gtgctgccgt	tgcttcccct	gctgcaggga	gagcggcaag	120

```

agcaacgtgg gcacttctgg agaccacgac gactctgcta tgaagacact caggagcaag 180
atgggcaagt ggtgccgcca ctgcttcccc tgctgcaggg ggagtggcaa gagcaacgtg 240
ggcgcttctg gagaccacga cgactctgct atgaagacac tcaggaacaa gatgggcaag 300
tgggtgctgcc actgcttccc ctgctgcagg gggagcggca agagcaaggt gggcgcttgg 360
ggagactacg atgacagtgc cttcatggag ccaggtacc acgtccgtgg agaagatctg 420
gacaagctcc acagagctgc ctggtggggg aaagtcccc gaaaggatct catcgtcatg 480
ctcagggaca ctgacgtgaa caagaaggac aagcaaaaga ggactgctct acatctggcc 540
tctgccaatg ggaattcaga agtagtaaaa ctctgctgg acagacgatg tcaacttaat 600
gtccttgaca acaaaaagag gacagctctg ataaaggccg tacaatgcca ggaagatgaa 660
tgtgcgttaa tgttgctgga acatggcact gatccaaata ttccagatga gtatggaaat 720
accactctgc actacgctat ctataatgaa gataaattaa tggccaaagc actgctctta 780
tatggtgctg atatcgaatc aaaaaacaag catggcctca caccactgtt acttgggtgta 840
catgagcaaa aacagcaagt cgtgaaatct ttaatcaaga aaaaagcgaa tttaaatgca 900
ctggatagat atggaaggac tgctctcata cttgctgtat gttgtggatc agcaagtata 960
gtcagccttc tacttgagca aaatattgat gtatcttctc aagatctatc tggacagacg 1020
gccagagagt atgctgtttc tagtcatcat catgtaattt gccagttact ttctgactac 1080
aaagaaaaac agatgctaaa aatctcttct gaaaacagca atccagaaca agacttaaag 1140
ctgacatcag aggaagagtc acaaagggtc aaaggcagtg aaaatagcca gccagagaaa 1200
atgtctcaag aaccagaaat aaataaggat ggtgatagag aggttgaaga agaaatgaag 1260
aagcatgaaa gtaataatgt gggattacta gaaaacctga ctaatggtgt cactgctggc 1320
aatggtgata atggattaat tctcaaagg aagagcagaa cacctgaaaa tcagcaattt 1380
cctgacaacg aaagtgaaga gtatcacaga atttgcgaat tagtttctga ctacaaagaa 1440
aaacagatgc caaaatactc ttctgaaaac agcaaccagc aacaagactt aaagctgaca 1500
tcagaggaag agtcacaaag gcttgagggc agtgaaaatg gccagccaga gaaaagatct 1560
caagaaccag aaataaataa ggatggtgat agagagctag aaaattttat ggctatcgaa 1620
gaaatgaaga agcacggaag tactcatgtc ggattcccag aaaacctgac taatggtgcc 1680
actgctggca atggtgatga tggattaatt cctccaagga agagcagaac acctgaaagc 1740
cagcaatttc ctgacactga gaatgaagag tatcacagt acgaacaaaa tgatactcag 1800
aagcaatttt gtgaagaaca gaacactgga atattacacg atgagattct gattcatgaa 1860
gaaaagcaga tagaagtggg tgaaaaaatg aattctgagc tttctcttag ttgtaagaaa 1920
gaaaagaca tcttgcagta aaatagtacg ttgcgggaag aaattgccat gctaagactg 1980
gagctagaca caatgaaaca tcagagccag ctaaaaaaaa aaaaaaaaaa aaaaaaaaaa 2040

```

<210> 376

<211> 329

<212> PRT

<213> Homo sapien

<400> 376

```

Met Asp Ile Val Val Ser Gly Ser His Pro Leu Trp Val Asp Ser Phe
 1             5             10             15
Leu His Leu Ala Gly Ser Asp Leu Leu Ser Arg Ser Leu Met Ala Glu
      20             25             30
Glu Tyr Thr Ile Val His Ala Ser Phe Ile Ser Cys Ile Ser Ser Ser
      35             40             45
Leu Asp Gly Gln Gly Glu Arg Gln Glu Gln Arg Gly His Phe Trp Arg
      50             55             60
Pro Gln Arg Leu Leu Cys Glu Asp Ala Trp Glu Gln Glu Val Gln Val
65             70             75             80
Val Leu Pro Leu Leu Pro Leu Leu Gln Gly Ser Gly Lys Ser Asn Val
      85             90             95

```

Val Ala Trp Gly Asp Tyr Asp Asp Ser Ala Phe Met Asp Pro Arg Tyr
 100 105 110
 His Val His Gly Glu Asp Leu Asp Lys Leu His Arg Ala Ala Trp Trp
 115 120 125
 Gly Lys Val Pro Arg Lys Asp Leu Ile Val Met Leu Arg Asp Thr Asp
 130 135 140
 Val Asn Lys Arg Asp Lys Gln Lys Arg Thr Ala Leu His Leu Ala Ser
 145 150 155 160
 Ala Asn Gly Asn Ser Glu Val Val Lys Leu Val Leu Asp Arg Arg Cys
 165 170 175
 Gln Leu Asn Val Leu Asp Asn Lys Lys Arg Thr Ala Leu Thr Lys Ala
 180 185 190
 Val Gln Cys Gln Glu Asp Glu Cys Ala Leu Met Leu Leu Glu His Gly
 195 200 205
 Thr Asp Pro Asn Ile Pro Asp Glu Tyr Gly Asn Thr Thr Leu His Tyr
 210 215 220
 Ala Val Tyr Asn Glu Asp Lys Leu Met Ala Lys Ala Leu Leu Leu Tyr
 225 230 235 240
 Gly Ala Asp Ile Glu Ser Lys Asn Lys His Gly Leu Thr Pro Leu Leu
 245 250 255
 Leu Gly Ile His Glu Gln Lys Gln Gln Val Val Lys Phe Leu Ile Lys
 260 265 270
 Lys Lys Ala Asn Leu Asn Ala Leu Asp Arg Tyr Gly Arg Thr Ala Leu
 275 280 285
 Ile Leu Ala Val Cys Cys Gly Ser Ala Ser Ile Val Ser Pro Leu Leu
 290 295 300
 Glu Gln Asn Val Asp Val Ser Ser Gln Asp Leu Glu Arg Arg Pro Glu
 305 310 315 320
 Ser Met Leu Phe Leu Val Ile Ile Met
 325

<210> 377

<211> 148

<212> PRT

<213> Homo sapien

<220>

<221> VARIANT

<222> (1)...(148)

<223> Xaa = Any Amino Acid

<400> 377

Met Thr Xaa Pro Ser Trp Ser Pro Gly Thr Thr Ser Val Glu Lys Ile
 1 5 10 15
 Trp Thr Ser Ser Thr Glu Leu Pro Trp Trp Gly Lys Val Pro Arg Lys
 20 25 30
 Asp Leu Ile Val Met Leu Arg Asp Thr Asp Val Asn Lys Xaa Asp Lys
 35 40 45
 Gln Lys Arg Thr Ala Leu His Leu Ala Ser Ala Asn Gly Asn Ser Glu
 50 55 60
 Val Val Lys Leu Xaa Leu Asp Arg Arg Cys Gln Leu Asn Val Leu Asp


```

Asn Gly Asp Asn Gly Leu Ile Pro Gln Arg Lys Ser Arg Thr Pro Glu
1490                      1495                      1500
Asn Gln Gln Phe Pro Asp Asn Glu Ser Glu Glu Tyr His Arg Ile Cys
1505                      1510                      1515                      152
Glu Leu Val Ser Asp Tyr Lys Glu Lys Gln Met Pro Lys Tyr Ser Ser
                      1525                      1530                      1535
Glu Asn Ser Asn Pro Glu Gln Asp Leu Lys Leu Thr Ser Glu Glu Glu
                      1540                      1545                      1550
Ser Gln Arg Leu Glu Gly Ser Glu Asn Gly Gln Pro Glu Lys Arg Ser
                      1555                      1560                      1565
Gln Glu Pro Glu Ile Asn Lys Asp Gly Asp Arg Glu Leu Glu Asn Phe
1570                      1575                      1580
Met Ala Ile Glu Glu Met Lys Lys His Gly Ser Thr His Val Gly Phe
1585                      1590                      1595                      160
Pro Glu Asn Leu Thr Asn Gly Ala Thr Ala Gly Asn Gly Asp Asp Gly
                      1605                      1610                      1615
Leu Ile Pro Pro Arg Lys Ser Arg Thr Pro Glu Ser Gln Gln Phe Pro
                      1620                      1625                      1630
Asp Thr Glu Asn Glu Glu Tyr His Ser Asp Glu Gln Asn Asp Thr Gln
                      1635                      1640                      1645
Lys Gln Phe Cys Glu Glu Gln Asn Thr Gly Ile Leu His Asp Glu Ile
1650                      1655                      1660
Leu Ile His Glu Glu Lys Gln Ile Glu Val Val Glu Lys Met Asn Ser
1665                      1670                      1675                      168
Glu Leu Ser Leu Ser Cys Lys Lys Glu Lys Asp Ile Leu His Glu Asn
                      1685                      1690                      1695
Ser Thr Leu Arg Glu Glu Ile Ala Met Leu Arg Leu Glu Leu Asp Thr
                      1700                      1705                      1710
Met Lys His Gln Ser Gln Leu
1715

```

<210> 379

<211> 656

<212> PRT

<213> Homo sapien

<400> 379

```

Met Val Val Glu Val Asp Ser Met Pro Ala Ala Ser Ser Val Lys Lys
1      5      10      15
Pro Phe Gly Leu Arg Ser Lys Met Gly Lys Trp Cys Cys Arg Cys Phe
20     25     30
Pro Cys Cys Arg Glu Ser Gly Lys Ser Asn Val Gly Thr Ser Gly Asp
35     40     45
His Asp Asp Ser Ala Met Lys Thr Leu Arg Ser Lys Met Gly Lys Trp
50     55     60
Cys Arg His Cys Phe Pro Cys Cys Arg Gly Ser Gly Lys Ser Asn Val
65     70     75     80
Gly Ala Ser Gly Asp His Asp Asp Ser Ala Met Lys Thr Leu Arg Asn
85     90     95
Lys Met Gly Lys Trp Cys Cys His Cys Phe Pro Cys Cys Arg Gly Ser
100    105    110

```


		515					520					525				
Lys	His	Gly	Ser	Thr	His	Val	Gly	Phe	Pro	Glu	Asn	Leu	Thr	Asn	Gly	
	530					535					540					
Ala	Thr	Ala	Gly	Asn	Gly	Asp	Asp	Gly	Leu	Ile	Pro	Pro	Arg	Lys	Ser	
545					550					555					560	
Arg	Thr	Pro	Glu	Ser	Gln	Gln	Phe	Pro	Asp	Thr	Glu	Asn	Glu	Glu	Tyr	
				565					570					575		
His	Ser	Asp	Glu	Gln	Asn	Asp	Thr	Gln	Lys	Gln	Phe	Cys	Glu	Glu	Gln	
			580					585				590				
Asn	Thr	Gly	Ile	Leu	His	Asp	Glu	Ile	Leu	Ile	His	Glu	Glu	Lys	Gln	
		595					600					605				
Ile	Glu	Val	Val	Glu	Lys	Met	Asn	Ser	Glu	Leu	Ser	Leu	Ser	Cys	Lys	
	610					615					620					
Lys	Glu	Lys	Asp	Ile	Leu	His	Glu	Asn	Ser	Thr	Leu	Arg	Glu	Glu	Ile	
625					630					635					640	
Ala	Met	Leu	Arg	Leu	Glu	Leu	Asp	Thr	Met	Lys	His	Gln	Ser	Gln	Leu	
				645					650					655		

```
<210> 380
<211> 671
<212> PRT
<213> Homo sapien
```

	<400>			380														
Met	Val	Val	Glu	Val	Asp	Ser	Met	Pro	Ala	Ala	Ser	Ser	Val	Lys	Lys			
1				5					10					15				
Pro	Phe	Gly	Leu	Arg	Ser	Lys	Met	Gly	Lys	Trp	Cys	Cys	Arg	Cys	Phe			
			20					25					30					
Pro	Cys	Cys	Arg	Glu	Ser	Gly	Lys	Ser	Asn	Val	Gly	Thr	Ser	Gly	Asp			
		35					40					45						
His	Asp	Asp	Ser	Ala	Met	Lys	Thr	Leu	Arg	Ser	Lys	Met	Gly	Lys	Trp			
	50					55					60							
Cys	Arg	His	Cys	Phe	Pro	Cys	Cys	Arg	Gly	Ser	Gly	Lys	Ser	Asn	Val			
65					70					75				80				
Gly	Ala	Ser	Gly	Asp	His	Asp	Asp	Ser	Ala	Met	Lys	Thr	Leu	Arg	Asn			
				85					90					95				
Lys	Met	Gly	Lys	Trp	Cys	Cys	His	Cys	Phe	Pro	Cys	Cys	Arg	Gly	Ser			
			100					105					110					
Gly	Lys	Ser	Lys	Val	Gly	Ala	Trp	Gly	Asp	Tyr	Asp	Asp	Ser	Ala	Phe			
		115					120					125						
Met	Glu	Pro	Arg	Tyr	His	Val	Arg	Gly	Glu	Asp	Leu	Asp	Lys	Leu	His			
	130					135					140							
Arg	Ala	Ala	Trp	Trp	Gly	Lys	Val	Pro	Arg	Lys	Asp	Leu	Ile	Val	Met			
145					150					155					160			
Leu	Arg	Asp	Thr	Asp	Val	Asn	Lys	Lys	Asp	Lys	Gln	Lys	Arg	Thr	Ala			
				165					170					175				
Leu	His	Leu	Ala	Ser	Ala	Asn	Gly	Asn	Ser	Glu	Val	Val	Lys	Leu	Leu			
			180					185					190					
Leu	Asp	Arg	Arg	Cys	Gln	Leu	Asn	Val	Leu	Asp	Asn	Lys	Lys	Arg	Thr			
		195					200					205						
Ala	Leu	Ile	Lys	Ala	Val	Gln	Cys	Gln	Glu	Asp	Glu	Cys	Ala	Leu	Met			

Glu Val Val Glu Lys Met Asn Ser Glu Leu Ser Leu Ser Cys Lys Lys
 625 630 635 640
 Glu Lys Asp Ile Leu His Glu Asn Ser Thr Leu Arg Glu Glu Ile Ala
 645 650 655
 Met Leu Arg Leu Glu Leu Asp Thr Met Lys His Gln Ser Gln Leu
 660 665 670

<210> 381
 <211> 251
 <212> DNA
 <213> Homo sapien

<400> 381
 ggagaagcgt ctgctggggc aggaaggggt ttccctgccc tctcacctgt cctcaccac 60
 ggtaacatgc ttcccctaag ggtatcccaa cccagggggc tcaccatgac ctctgagggg 120
 ccaatatccc aggagaagca ttggggaggt gggggcaggt gaaggaccca ggactcacac 180
 atcctggggc tccaaggcag aggagaggggt cctcaagaag gtcaggagga aaatccgtaa 240
 caagcagtca g 251

<210> 382
 <211> 3279
 <212> DNA
 <213> Homo sapiens

<400> 382
 ctctcctgcag ccccatgct ggtgaggggc acgggcagga acagtggacc caacatggaa 60
 atgctggagg gtgtcaggaa gtgatcgggc tctggggcag ggaggagggg tggggagtgt 120
 cactgggagg ggacatcctg cagaaggtag gagtgcagca acaccgctg caggggaggg 180
 gagagccctg cggcacctgg gggagcagag ggagcagcac ctgcccaggc ctgggaggag 240
 gggcctggag ggcgtgagga ggagcgaggg ggctgcattg ctggagtgcg ggatcagggg 300
 cagggcgcgga gatggcctca cacaggggaag agaggggccc tctgcaggg cctcacctgg 360
 gccacaggag gacactgctt ttctctgcag gagtgcaggag ctgtggatgg tgctggacag 420
 aagaaggaca gggcctggct caggtgtcca gaggtgtcg ctggcttccc tttgggatca 480
 gactgcaggg agggagggcg gcaggggtgt ggggggagtg acgatgagga tgacctgggg 540
 gtggctccag gccttgcacc tgctggggc ctcaccagc ctccctcaca gtctcctggc 600
 cctcagtcct tcccctccac tccatcctcc atctggcctc agtgggtcat tctgatcact 660
 gaactgacca taccagccc tgcccacggc cctccatggc tcccacatgc cctggagagg 720
 ggacatctag tcagagagta gtctgaaga ggtggcctct gcgatgtgcc tgtgggggca 780
 gcatcctgca gatggtcccg gccctcatcc tgcctgacct tctgcaggga ctgtcctcct 840
 ggaccttgcc ccttgtgcag gagctggacc ctgaagtccc ctcccatag gccaaactg 900
 gagccttggt cctctgtgtg gactccctgc ccatattctt gtgggagtgg gttctggaga 960
 catttctgtc tgttctgcag agctgggaat tgctctcagt catctgcctg cgcggttctg 1020
 agagatggag ttgcctaggc agttattggg gccaatcttt ctactgtgt ctctcctcct 1080
 ttacccttag ggtgattctg ggggtccact tgtctgtaat ggtgtgcttc aaggtatcac 1140
 atcatggggc cctgagccat gtgccctgcc tgaaaagcct gctgtgtaca ccaaggtggt 1200
 gcattaccgg aagtggatca aggacaccat cgcagccaac ccctgagtgc ccctgtccca 1260
 cccctacctc tagtaaatct aagtccacct caggttctgg catcacttgg ctttctgga 1320
 tgctggacac ctgaagcttg gaactcacct ggccgaagct cgagcctcct gactcctact 1380
 gacctgtgct ttctgggtgt gagtccaggg ctgctaggaa aaggaatggg cagacacagg 1440
 tgtatgccaa tgtttctgaa atgggtataa ttctgcctc tccttcggaa cactggctgt 1500
 ctctgaagac ttctgcgtca gtttcagtga ggacacacac aaagacgtgg gtgacctgt 1560

tgtttgtggg gtgcagagat gggaggggtg gggccacccc tggaagagtg gacagtgaca 1620
 caaggtggac actctctaca gatcactgag gataagctgg agccacaatg catgaggcac 1680
 acacacagca aggttgacgc tgtaaacata gccacgctg tcttgggggc actgggaagc 1740
 ctagataagg ccgtgagcag aaagaagggg aggatcctcc tatgttggtg aaggagggac 1800
 tagggggaga aactgaaagc tgattaatta caggagggtt gttcagggtcc cccaaaccac 1860
 cgtcagattt gatgatttcc tagcaggact tacagaaata aagagctatc atgctgtggt 1920
 ttattatggt ttgttacatt gataggatac atactgaaat cagcaaacaa aacagatgta 1980
 tagattagag tgtggagaaa acagaggaaa acttgcagtt acgaagactg gcaacttggc 2040
 ttactaagt tttcagactg gcaggaagtc aaacctatta ggctgaggac cttgtggagt 2100
 gtagctgac cagctgatag aggaactagc cagggtggggg cctttccctt tggatggggg 2160
 gcataatccga cagttattct ctccaagtgg agacttacgg acagcatata attctccctg 2220
 caaggatgta tgataatatg tacaaagtaa ttccaactga ggaagctcac ctgacctta 2280
 gtgtccaggg tttttactgg ggggtctgtg gacgagtatg gactacttga ataattgacc 2340
 tgaagtcctc agacctgagg ttccctagag ttcaaacaga tacagcatgg tccagagtcc 2400
 cagatgtaca aaaacaggga ttcatcaca atcccatctt tagcatgaag ggtctggcat 2460
 ggcccaaggc cccaagtata tcaaggcact tgggcagAAC atgccaagga atcaaagtgc 2520
 atctcccagg agttattcaa ggggtgagccc tttacttggg atgtacaggc tttgagcagt 2580
 gcagggctgc tgagtcaacc ttttattgta caggggatga gggaaaggga gaggatgagg 2640
 aagccccctt ggggatttgg tttgggtctt tgatcagggt gtctatgggg ctatccctac 2700
 aaagaagaat ccagaaatag gggcacattg aggaatgata ctgagcccaa agagcattca 2760
 atcattgttt tatttgcctt cttttcacac cattgggtgag ggagggatta ccaccctggg 2820
 gttatgaaga tgggtgaaca cccacacat agcaccggag atatgagatc aacagtttct 2880
 tagccataga gattcacagc ccagagcagg aggacgctgc acaccatgca ggatgacatg 2940
 ggggatgcgc tcgggattgg tgtgaagaag caaggactgt tagaggcagg ctttatagta 3000
 acaagacggt ggggcaaact ctgatttccg tgggggaatg tcatggtctt gctttactaa 3060
 gttttgagac tggcaggtag tgaaactcat taggctgaga accttgtgga atgcagctga 3120
 cccagctgat agaggaaagta gccagggtggg agcctttccc agtgggtgtg ggacatatct 3180
 ggcaagattt tgtggcactc ctggttacag atactggggc agcaaataaa actgaatctt 3240
 gttttcagac cttaaaaaaa aaaaaaaaaa aaaagtttt 3279

<210> 383

<211> 155

<212> PRT

<213> Homo sapiens

<400> 383

Met Ala Gly Val Arg Asp Gln Gly Gln Gly Ala Arg Trp Pro His Thr
 5 10 15
 Gly Lys Arg Gly Pro Leu Leu Gln Gly Leu Thr Trp Ala Thr Gly Gly
 20 25 30
 His Cys Phe Ser Ser Glu Glu Ser Gly Ala Val Asp Gly Ala Gly Gln
 35 40 45
 Lys Lys Asp Arg Ala Trp Leu Arg Cys Pro Glu Ala Val Ala Gly Phe
 50 55 60
 Pro Leu Gly Ser Asp Cys Arg Glu Gly Gly Arg Gln Gly Cys Gly Gly
 65 70 75 80

Ser Asp Asp Glu Asp Asp Leu Gly Val Ala Pro Gly Leu Ala Pro Ala
 85 90 95

Trp Ala Leu Thr Gln Pro Pro Ser Gln Ser Pro Gly Pro Gln Ser Leu
 100 105 110

Pro Ser Thr Pro Ser Ser Ile Trp Pro Gln Trp Val Ile Leu Ile Thr
 115 120 125

Glu Leu Thr Ile Pro Ser Pro Ala His Gly Pro Pro Trp Leu Pro Asn
 130 135 140

Ala Leu Glu Arg Gly His Leu Val Arg Glu
 145 150

<210> 384

<211> 557

<212> DNA

<213> Homo sapiens

<400> 384

```
ggatcctcta gagcgccgc ctactactac taaattcgcg gccgcgtcga cgaagaagag 60
aaagatgtgt tttgttttgg actctctgtg gtcccttcca atgctgtggg tttccaacca 120
ggggaagggt cccttttgca ttgccaagtg ccataaccat gagcactact ctaccatggg 180
tctgcctcct ggccaagcag gctggtttgc aagaatgaaa tgaatgattc tacagctagg 240
acttaacctt gaaatggaaa gtcttgcaat cccatttgca ggatccgtct gtgcacatgc 300
ctctgtagag agcagcattc ccagggacct tggaaacagt tggcactgta aggtgcttgc 360
tccccaagac acatcctaaa aggtgttgta atgggtgaaaa cgtcttccct ctttattgcc 420
ccttcttatt tatgtgaaca actgtttgtc tttttttgta tcttttttaa actgtaaagt 480
tcaattgtga aaatgaatat catgcaaata aattatgcga tttttttttc aaagtaaaaa 540
aaaaaaaaaa aaaaaaaa                                     557
```

<210> 385

<211> 337

<212> DNA

<213> Homo sapiens

<400> 385

```
ttcccagggt atgtgcgagg gaagacacat ttactatcct tgatggggct gattccttta 60
gtttctctag cagcagatgg gttaggagga agtgacccaa gtggttgact cctatgtgca 120
tctcaaagcc atctgctgtc ttcgagtagc gacacatcat cactcctgca ttgttgatca 180
aaacgtggag gtgcttttcc tcagctaaga agcccttagc aaaagctcga atagacttag 240
tatcagacag gtccagtttc cgcaccaaca cctgctgggt ccctgtcgtg gtctggatct 300
ctttggccac caattccccc ttttccacat cccggca                                     337
```

<210> 386

<211> 300

<212> DNA

<213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(183)
 <223> n = A,T,C or G

<400> 408
 ggagctngcc ctcaattcct ccatntctat gttancatat ttaatgtctt ttgnnattaa 60
 tncttaacta gttaatcctt aaagggctan ntaatcetta actagtcctt ccattgtgag 120
 cattatcctt ccagtattcn ccttctnttt tatttactcc ttcttggtta cccatgtact 180
 ntt 183

<210> 409
 <211> 250
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(250)
 <223> n = A,T,C or G

<400> 409
 cccacgcatg ataagctctt tatttctgta agtcctgcta ggaaatcatc aaatctgacg 60
 gtggtttggg ggacctgaac aaacctcctg taattaatca gctttcagtt tctcccccta 120
 gtccctcctt caacaacata ggaggatcct ccccttcttt ctgctcacgg ccttatctag 180
 gcttcccagt gccccagga cagcgtgggc tatgtttaca gcgcntcctt gctggggggg 240
 ggccttatgc 250

<210> 410
 <211> 306
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(306)
 <223> n = A,T,C or G

<400> 410
 ggctggtttg caagaatgaa atgaatgatt ctacagctag gacttaacct tgaaatggaa 60
 agtcttgcaa tcccatattgc aggatccgtc tgtgcacatg cctctgtaga gagcagcatt 120
 cccagggacc ttggaaacag ttggcactgt aagggtgctt ctcccccaaga cacatcctaa 180
 aagggtgttg aatggtgaaa accgcttcct tctttattgc cccttcttat ttatgtgaac 240
 nactggttgg ctttttttgn atctttttta aactggaaag ttcaattgng aaaatgaata 300
 tcntgc 306

<210> 411
 <211> 261
 <212> DNA
 <213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(261)

<223> n = A,T,C or G

<400> 411

```
agagatattn cttaggtnaa agttcataga gttcccatga actatatgac tggccacaca 60
ggatcttttg tattaagga ttctgagatt ttgcttgagc aggattagat aaggctgttc 120
tttaaagtgc tgaaatggaa cagatttcaa aaaaaaaccc cacaatctag ggtgggaaca 180
aggaaggaaa gatgtgaata ggctgatggg caaaaaacca atttaccat cagttccagc 240
cttctctcaa ggngaggcaa a 261
```

<210> 412

<211> 241

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(241)

<223> n = A,T,C or G

<400> 412

```
gttcaatggt acctgacatt tctacaacac cccactcacc gatgtattcg ttgcccagtg 60
ggaacatacc agcctgaatt tggaaaaaat aattgtgttt cttgcccagg aaatactacg 120
actgactttg atggctccac aaacataacc cagtgtaaaa acagaagatg tggaggggag 180
ctgggagatt tctactgggt cattgaattc ccaaactacc cangcaatta ccagccaac 240
a 241
```

<210> 413

<211> 231

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(231)

<223> n = A,T,C or G

<400> 413

```
aactcttaca atccaagtga ctcatctgtg tgcttgaatc ctttccactg tctcatctcc 60
ctcatccaag tttctagtag cttctctttg ttgtgaagga taatcaaact gaacaacaaa 120
aagtttactc tcctcatttg gaacctaaaa actctcttct tcctgggtct gagggctcca 180
agaatccttg aatcanttct cagatcattg gggacaccan atcaggaacc t 231
```

<210> 414

<211> 234

<212> DNA

<213> Homo sapiens

<400> 414


```

gccaaactcac ccagctgggc atggagcagc attatgaact tggagagtat ataagaaaga 300
gatatagaaa attcttgaat gagtcctata aacatgaaca ggtttatatt cgaagcacag 360
acgttgaccg gactttgatg aagtgcctatg acaaacctgg caagcccg 408

```

<210> 421

<211> 352

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(352)

<223> n = A,T,C or G

<400> 421

```

gctcaaaaat ctttttactg atnggcatgg ctacacaatc attgactatt acggaggcca 60
gaggagaatg aggcctggcc tgggagccct gtgcctacta naagcacatt agattatcca 120
ttcactgaca gaacaggtct tttttgggtc cttcttctcc accacnatac acttgcagtc 180
ctccttcttg aagattcttt ggcagttgtc tttgtcataa cccacaggtg tagaaacaag 240
ggtgcaacat gaaatttctg tttcgtagca agtgcctgtc tcacaagttg gcangtctgc 300
cactccgagt ttattgggtg tttgtttcct ttgagatcca tgcatttctc gg 352

```

<210> 422

<211> 337

<212> DNA

<213> Homo sapiens

<400> 422

```

atgccaccat gctggcaatg cagcggggcg tcgaaggcct gcatatccag cccaagctgg 60
cgatgatcga cggcaaccgt tgcccgaggt tgccgatgcc agccgaagcg gtggtcaagg 120
gcgatagcaa ggtgccggcg atcgcgggcg cgtcaatcct ggccaaggtc agccgtgatc 180
gtgaaatggc agctgtcgaa ttgatctacc cgggttatgg catcgggcggg cataagggtc 240
atccgacacc ggtgcacctg gaagccttgc agcggtctggg gccgacgccg attcaccgac 300
gcttcttccg ccggtacggc tggcctatga aaattat 337

```

<210> 423

<211> 310

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(310)

<223> n = A,T,C or G

<400> 423

```

gctcaaaaat ctttttactg atatggcatg gctacacaat cattgactat tagaggccag 60
aggagaatga ggccctggcct gggagccctg tgccctactan aagcncatta gattatccat 120
tcactgacag aacaggtctt ttttgggtcc ttcttctcca ccacgatata cttgcagtc 180
tccttcttga agattctttg gcagttgtct ttgtcataac ccacaggtgt anaaacaagg 240
gtgcaacatg aaatttctgt ttcgtagcaa gtgcctgtct cacagttgtc aagtctgccc 300

```


atacactcat atactcgtgg gcttagaggg cacagcagat gtcattgggc tactgcctga 540
gtcccgcgtgg tcccatccca ggaccttcca tcggcgagta cctgggagcc cgtgct 596

<210> 427

<211> 107

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(107)

<223> n = A,T,C or G

<400> 427

gaagaattca agttaggttt attcaaaggg cttacngaga atcctanacc caggncaccag 60
cccgggagca gccttanaga gtcctgttt gactgcccggt ctcagng 107

<210> 428

<211> 38

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(38)

<223> n = A,T,C or G

<400> 428

gaacttcna anaangactt tattcactat ttacatt 38

<210> 429

<211> 544

<212> DNA

<213> Homo sapiens

<400> 429

ctttgctgga cggaataaaa gtggacgcaa gcatgacctc ctgatgaggg cgctgcattt 60
attgaagagc ggctgcagcc ctgoggttca gattaaaatc cgagaattgt atagacgccg 120
atatccacga actcctgaag gactttctga tttatccaca atcaaatacat cggtttttcag 180
tttggtggtt ggctcatcac ctgtagaacc tgacttggcc gtggctggaa tccactcgtt 240
gccttccact tcagttacac ctcactcacc atcctctcct gttgggttctg tgctgcttca 300
agatactaag cccacatttg agatgcagca gccatctccc ccaattcctc ctgtccatcc 360
tgatgtgcag ttaaaaaatc tgccctttta tgatgtcctt gatgttctca tcaagcccac 420
gagtttagtt caaagcagta ttcagcgatt tcaagagaag ttttttattt ttgctttgac 480
acctcaacaa gtttagagaga tatgcatatc cagggatttt ttgccaggtg gtaggagaga 540
ttat 544

<210> 430

<211> 507

<212> DNA

<213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(507)
 <223> n = A,T,C or G

<400> 430
 cttatcncaa tggggctccc aaacttggct gtgcagtgga aactccgggg gaattttgaa 60
 gaacactgac acccatcttc caccgagaca ctctgattta attgggctgc agtgagaaca 120
 gagcatcaat ttaaaaaagct gcccgagaatg ttntcctggg cagcgttggtg atctttgccn 180
 ccttcgtgac tttatgcaat gcatcatgct atttcatacc taatgaggga gttccaggag 240
 attcaaccag gatgtttcta cncctgtggg ttatgacaaa gacaactgcc aaagaatntt 300
 caagaaggag gactgcaagt atatcgtggt ggagaagaag gacccaaaaa agacctgttc 360
 tgtcagtga tggataatct aatgtgcttc tagtaggcac agggctccca ggccaggcct 420
 cattctctc tggcctctaa tagtcaatga ttgtgtagcc atgcctatca gtaaaaagat 480
 ttttgagcaa aaaaaaaaaa aaaaaaa 507

<210> 431
 <211> 392
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(392)
 <223> n = A,T,C or G

<400> 431
 gaaaattcag aatggataaa aacaaatgaa gtacaaaata tttcagattt acatagcgat 60
 aaacaagaaa gcacttatca ggaggactta caaatggaag tacactctan aaccatcatc 120
 tatcatggct aaatgtgaga ttagcacagc tgtattattt gtacattgca aacacctaga 180
 aagagatggg aaacaaaatc ccaggagttt tgtgtgtgga gtccctgggtt ttccaacaga 240
 catcattcca gcattctgag attagggnga ttggggatca ttctggagtt ggaatgttca 300
 acaaaaagtga tgttgttagg taaaatgtac aacttctgga tctatgcaga cattgaaggt 360
 gcaatgagtc tggcttttac tctgctgttt ct 392

<210> 432
 <211> 387
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(387)
 <223> n = A,T,C or G

<400> 432
 ggtatccnta cataatcaaa tatagctgta gtacatgttt tcattggngt agattaccac 60
 aaatgcaagg caacatgtgt agatctcttg tcttattctt ttgtctataa tactgtattg 120
 ngtagtccaa gctctcggn a gtccagccac tngnaaacat gctcccttta gattaacctc 180
 gtggacnctn ttgttgnatt gtctgaactg tagngccctg tattttgctt ctgtctgnga 240

```

attctgttgc ttctggggca tttccttgng atgcagagga ccaccacaca gatgacagca 300
atctgaattg ntccaatcac agctgcgatt aagacatact gaaatcgtac aggaccggga 360
acaacgtata gaacactgga gtccttt 387

```

<210> 433

<211> 281

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(281)

<223> n = A,T,C or G

<400> 433

```

ttcaactagc anagaanact gcttcagggg gtgtaaaatg aaaggcttcc acgcagttat 60
ctgattaaag aacactaaga gagggacaag gctagaagcc gcaggatgtc tacactatag 120
caggcnctat ttgggttggc tggaggagct gtggaaaaca tggagagatt ggcgctggag 180
atcgccgtgg ctattcctcn ttgntattac accagngagg ntctctgtnt gccactgggt 240
tnnaaaaccg ntatacaata atgatagaat aggacacaca t 281

```

<210> 434

<211> 484

<212> DNA

<213> Homo sapiens

<400> 434

```

ttttaaaata agcatttagt gctcagtcct tactgagtac tctttctctc cctcctctctg 60
aatttaattc tttcaacttg caatttgcaa ggattacaca tttcactgtg atgtatattg 120
tgttgcaaaa aaaaaaaagt gtctttgttt aaaattactt ggtttgtaga tccatcttgc 180
tttttcccca ttggaactag tcattaacct atctctgaac tggtagaaaa acatctgaag 240
agctagtcta tcagcatctg acaggtgaat tggatggttc tcagaacctt ttcaccaga 300
cagcctgttt ctatcctgtt taataaatta gtttggttct tctacatgca taacaaacct 360
tgctccaatc tgtcacataa aagtctgtga cttgaagttt agtcagcacc cccaccaaac 420
tttatttttc tatgtgtttt ttgcaacata tgagtgtttt gaaaataaag taccatgtc 480
ttta 484

```

<210> 435

<211> 424

<212> DNA

<213> Homo sapiens

<400> 435

```

gcgccgctca gagcaggtea ctttctgcct tccacgtcct ccttcaagga agcccatgt 60
gggtagcttt caatatcgca ggttcttact cctctgcctc tataagctca aaccaccaa 120
cgatcgggca agtaaacccc ctccctcgcc gacttcggaa ctggcgagag ttcagcgag 180
atgggcctgt ggggaggggg caagatagat gagggggagc ggcatgggtc ggggtgacct 240
cttgagagaga ggaaaaaggc cacaagaggg gctgccaccg ccactaacgg agatggcct 300
ggtagagacc tttgggggtc tggaaacctc ggactcccca tgctctaact cccacactct 360
gctatcagaa acttaaacctt gaggattttc tctgtttttc actcgcaata aattcagagc 420
aaac 424

```

<210> 436
 <211> 667
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(667)
 <223> n = A,T,C or G

<400> 436
 accttgggaa nactctcaca atataaaggg tcgtagactt tactccaaat tccaaaaagg 60
 tcctggccat gtaatcctga aagttttccc aaggtagcta taaaatcctt ataaggggtgc 120
 agcctcttct ggaattcctc tgatttcaaa gtctcactct caagttcttg aaaacgaggg 180
 cagttcctga aaggcaggta tagcaactga tcttcagaaa gaggaactgt gtgcaccggg 240
 atgggctgcc agagtaggat aggattccag atgctgacac cttctggggg aaacaggggt 300
 gccaggtttg tcatagcact catcaaagtc cggtcacgt ctgtgcttcg aatataaacc 360
 tgttcatgtt tataggactc attcaagaat tttctatata tctttcttat atactctcca 420
 agttcataat gctgctccat gccagctgg gtgagttggc caaatccttg tggccatgag 480
 gattccttta tggggtcagt gggaaagggt tcaatgggac ttcgggtctcc atgccgaaac 540
 accaaagtca caaacttcaa ctcttgggt agtacacttc ggtctagcca gaaaaaagc 600
 agaaacaaga agccaagggt aaggcttgct gccctgccag gaggaggggt gcagctctca 660
 tgttgag 667

<210> 437
 <211> 693
 <212> DNA
 <213> Homo sapiens

<400> 437
 ctacgtctca accctcattt ttaggtaagg aatcttaagt ccaaagatat taagtgactc 60
 acacagccag gtaaggaaaag ctggattggc acactaggac tctaccatac cgggttttgt 120
 taaagctcag gttaggaggc tgataagctt ggaagggaact tcagacagct ttttcagatc 180
 ataaaagata attcttagcc catgttcttc tcagagcag acctgaaatg acagcacagc 240
 aggtactcct ctattttcac ccctcttgct tctactctct ggcagtcaga cctgtgggag 300
 gccatgggag aaagcagctc tctggatgtt tgtacagatc atggactatt ctctgtggac 360
 catttctcca ggttacccta ggtgtcacta ttgggggggac agccagcatc tttagctttc 420
 atttgagttt ctgtctgtct tcagtagagg aaacttttgc tcttcacact tcacatctga 480
 acacctaact gctgttgctc ctgaggtggg gaaagacaga tatagagctt acagtattta 540
 tcctatttct aggcactgag ggctgtgggg taccttgtgg tgccaaaaca gatcctgttt 600
 taaggacatg ttgcttcaga gatgtctgta actatctggg ggctctgttg gctctttacc 660
 ctgcatcatg tgctctcttg gctgaaaatg acc 693

<210> 438
 <211> 360
 <212> DNA
 <213> Homo sapiens

<400> 438
 ctgcttatca caatgaatgt tctcctgggc agcgttgtga tctttgccac cttcgtgact 60


```
<210> 439
<211> 431
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> (1)...(431)
<223> n = A,T,C or G
```

```

<210> 440
<211> 523
<212> DNA
<213> Homo sapiens

<400> 440
agagataaag cttaggtcaa agttcataga gttcccatga actatatgac tggccacaca 60
ggatccttttg tatttaagga ttctgagatt ttgcttgagc aggattagat aaggctgttc 120
tttaaatgtc tgaaatggaa cagatttcaa aaaaaaaccc cacaatctag ggtgggaaca 180
aggaaggaaa gatgtgaata ggctgatggg caaaaaacca atttaccat cagttccagc 240
cttctctcaa ggagaggcaa agaaaggaga tacagtggag acatctggaa agttttctcc 300
actggaaaac tgctactatc tgtttttata tttctgttaa aatatatgag gctacagaac 360
taaaaattaa aacctctttg tgtcccttgg tcctggaaca tttatgttcc ttttaaagaa 420
acaaaaatca aactttacag aaagatttga tgtatgtaat acatatagca gctcttgaag 480
tatatatatc atagcaaata agtcacttga tgagaacaag cta 523

```

```
<210> 441
<211> 430
<212> DNA
<213> Homo sapiens

<400> 441
gttcctccta actcctgcc aaacagctc tctcaacat gagagctgca cccctcctcc 60
tgccaggggc agcaagcctt agccttggct tcttgtttct gctttttttc tggctagacc 120
gaagtgtact agccaaggag ttgaagtttg tgactttqgt qtttcggcat qqagaccqaa 180
```

```
<210> 442
<211> 362
<212> DNA
<213> Homo sapiens
```

```
<210> 443
<211> 624
<212> DNA
<213> Homo sapiens
```

<400>	443						
tttttttttt	gcaacacaat	atacatcaca	gtgaaatgtg	taatccttgc	aaattgcaag	60	
ttgaaagaat	taaattcaga	ggaggggaga	gaaagagtac	tcagtaggga	ctgagcacta	120	
aatgcttatt	ttaaagaaa	tgtaaagagc	agaaagcaat	tcaggctacc	ctgccttttg	180	
tgctgggctag	tactccggtc	ggtgtcagca	gcacgtggca	ttgaacattg	caatgtggag	240	
cccaaaccac	agaaaatggg	gtgaaattgg	ccaactttct	attaacttgg	cttcctgttt	300	
tataaaatat	tgtgaataat	atcacctact	tcaaagggca	gttatgaggc	ttaaatgaac	360	
taacgcctac	aaaacactta	aacatagata	acataggtgc	aagtactatg	tatctggtac	420	
atggtaaaca	tccttattat	taaagtcac	gctaaaatga	atgtgtgtgc	atatgtcta	480	
agtacagaga	gagggcactt	aaaccaacta	agggcctgga	gggaagggtt	cctggaaaga	540	
ngatgctttgt	gctgggtcca	aatcttggtc	tactatgacc	ttggccaaat	tatttaaact	600	
ttgtccctat	ctqcta	aaacagatc				624	

```
<220>  
<221> misc_feature  
<222> (1)...(425)
```

<223> n = A,T,C or G

<400> 444

```
gcacatcatt nntcttgcatt tctttgagaa taagaagatc agtaaatagt tcagaagtgg 60
gaagctttgt ccaggcctgt gtgtgaaccc aatgttttgc ttagaaatag aacaagtaag 120
ttcattgcta tagcataaca caaaatttgc ataagtgggtg gtcagcaaatt ccttgaatgc 180
tgcttaaatgt gagaggttgg taaaatcctt tgtgcaaacac tctaactccc tgaatgtttt 240
gctgtgctgg gacctgtgca tgccagacaa ggccaagctg gctgaaagag caaccagcca 300
cctctgcaat ctgccacctc ctgctggcag gatttgtttt tgcacacctg gaagagccaa 360
ggaggcacca gggcataagt gagtagactt atgggtcgacg cggccgcgaa tttagtagta 420
gtaga 425
```

<210> 445

<211> 414

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(414)

<223> n = A,T,C or G

<400> 445

```
catgtttatg nttttggatt actttgggca cctagtgttt ctaaatcgtc tatcattctt 60
ttctgttttt caaaagcaga gatggccaga gtctcaacaa actgtatctt caagtctttg 120
tgaaattctt tgcattgtggc agattatttg atgtagtctt ctttaactag catataaatc 180
tgggtgtgtt cagataaatg aacagcaaaa tgtggtggaa ttaccatttg gaacattgtg 240
aatgaaaaat tgtgtctcta gattatgtaa caaataacta tttcctaacc attgatcttt 300
ggatttttat aatcctactc acaaatgact aggtctctcc tcttgtattt tgaagcagt 360
tgggtgctgg attgataaaa aaaaaaaaaa tgcacgcggc cgcgaattta gtag 414
```

<210> 446

<211> 631

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(631)

<223> n = A,T,C or G

<400> 446

```
acaaattaga anaaagtgcc agagaacacc acataccttg tccggaacat tacaatggct 60
tctgcatgca tgggaagtgt gagcattcta tcaatatgca ggagccatct tgcagggtgtg 120
atgctggtta tactggacaa cactgtgaaa aaaaggacta cagtgttcta tacgttgttc 180
ccggtcctgt acgatttcag tatgtcttaa tcgcagctgt gattggaaca attcagattg 240
ctgtcatctg tgtggtgggtc ctctgcatca caagggccaa actttaggta atagcattgg 300
actgagattt gtaaaacttc caacctcca ggaaatgccc cagaagcaac agaattcaca 360
gacagaagca aaatacaggg cactacagtt cagacaatac aacaagagcg tccacgaggt 420
taatctaaag ggagcatgtt tcacagtggc tggactaccg agagcttgga ctacacaata 480
cagtattata gacaaaagaa taagacaaga gatctacaca tgttgccttg catttgggt 540
```

```
aatctacacc aatgaaaaca tgtactacag ctatatattga ttatgtatgg atatatttga 600
aatagtatac attgtcttga tgttttttct g 631
```

<210> 447

<211> 585

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(585)

<223> n = A,T,C or G

<400> 447

```
ccttgggaaa antntcacaa tataaagggt cgtagacttt actccaaatt ccaaaaagggt 60
cctggccatg taatcctgaa agttttccca aggtagctat aaaatcctta taagggtgca 120
gcctcttctg gaattcctct gatttcaaag tctcactctc aagttcttga aaacgagggc 180
agttcctgaa aggcaggtat agcaactgat cttcagaaag aggaactgtg tgcaccggga 240
tggtctgcca gagtaggata ggattccaga tgcagacacc ttctggggga aacagggctg 300
ccaggtttgt catagcactc atcaaagtcc ggtcaacgtc tgtgcttcga atataaacct 360
gttcataatg ctgctccatg cccagctggg tgagttggcc aaatccttgt ggccatgagg 480
attcctttat ggggtcagtg ggaaagggtg caatgggact tcggtctcca tgccgaaaca 540
ccaaagtcac aaacttcaac tccttggcta gtacacttcg gtcta 585
```

<210> 448

<211> 93

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(93)

<223> n = A,T,C or G

<400> 448

```
tgctcgtggg tcattctgan nnccgaactg acntgtccag ccctgccgan gggccnccat 60
ggctccctag tgccttgag agganggggc tag 93
```

<210> 449

<211> 706

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(706)

<223> n = A,T,C or G

<400> 449

```
ccaagttcat gctntgtgct ggacgctgga cagggggcaa aagcnnttgc tcgtgggtca 60
```

```

ttctgancac cgaactgacc atgccagccc tgccgatggc cctccatggc tccctagtgc 120
cctggagagg aggtgtctag tcagagagta gtccctggaag gtggcctctg ngaggagcca 180
cggggacagc atcctgcaga tggtcgggcg cgtcccattc gccattcagg ctgcgcaact 240
gttgggaagg gcgatcggtg cgggcctctt cgctattacg ccagctggcg aaaggggggat 300
gtgctgcaag gcgattaagt tgggtaacgc cagggttttc ccagtcncga cgttgtaaaa 360
cgacggccag tgaattgaat ttaggtgacn ctatagaaga gctatgacgt cgcattgcacg 420
cgtacgtaag cttggatcct ctagagcggc cgcctactac tactaaattc gcggccgcgt 480
cgacgtggga tccnactga gagagtggag agtgacatgt gctggacnct gtccatgaag 540
cactgagcag aagctggagg cacaacgcnc cagacactca cagctactca ggaggctgag 600
aacaggttga acctgggagg tggaggttgc aatgagctga gatcaggccn ctgcncacca 660
gcatggatga cagagtgaaa ctccatctta aaaaaaaaaa aaaaaa 706

```

<210> 450

<211> 493

<212> DNA

<213> Homo sapiens

<400> 450

```

gagacggagt gtcactctgt tgcccaggct ggagtgcagc aagacactgt ctaagaaaaa 60
acagttttaa aaggtaaaac aacataaaaa gaaatatact atagtggaaa taagagagtc 120
aaatgaggct gagaacttta caaagggatc ttacagacat gtcgccaata tcaactgcag 180
agcctaagta taagaacaac ctttggggag aaaccatcat ttgacagtga ggtacaattc 240
caagtcaggt agtgaaatgg gtggaattaa actcaaatta atcctgccag ctgaaacgca 300
agagacactg tcagagagtt aaaaagttag ttctatccat gaggtgattc cacagtcttc 360
tcaagtcaac acatctgtga actcacagac caagttctta aaccactgtt caaactctgc 420
tacacatcag aatcacctgg agagctttac aaactcccat tgccgagggt cgacgcggcc 480
gcgaatttag tag 493

```

<210> 451

<211> 501

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(501)

<223> n = A,T,C or G

<400> 451

```

gggcgcgtcc cattcgccat tcaggctgcg caactgttgg gaagggcgat cgggtgcgggc 60
ctcttcgcta ttacgccagc tggcgaaagg gggatgtgct gcaaggcgat taagttgggt 120
aacgccaggg ttttcccagt cncgacgttg taaaacgacg gccagtgaat tgaatttagg 180
tgacnctata gaagagctat gacgtcgcat gcacgcgtac gtaagcttgg atcctctaga 240
gcggccgcct actactacta aattcgcggc cgcgtcgacg tgggatccnc actgagagag 300
tggagagtga catgtgctgg acnctgtcca tgaagcactg agcagaagct ggaggcacia 360
cgcncagac actcacagct actcaggagg ctgagaacag gttgaacctg ggaggtggag 420
gttgcaatga gctgagatca ggccnctgcn ccccgacatg gatgacagag tgaaactcca 480
tcttaaaaaa aaaaaaaaaa a 501

```

<210> 452

<211> 51

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1) ... (51)

<223> n = A,T,C or G

<400> 452

agacgggtttc accnttacaa cnccttttag gatgggnntt ggggagcaag c 51

<210> 453

<211> 317

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1) ... (317)

<223> n = A,T,C or G

<400> 453

tacatcttgc tttttcccca ttggaactag tcattaaccc atctctgaac tggtagaaaa 60
 acatctgaag agctagtcta tcagcatctg gcaagtgaat tggatgggtc tcagaaccat 120
 ttcaccana cagcctgttt ctatcctgtt taataaatta gtttgggttc tctacatgca 180
 taacaaaccc tgctccaatc tgtcacataa aagtctgtga cttgaagttt antcagcacc 240
 cccaccaaac tttatttttc tatgtgtttt ttgcaacata tgagtgtttt gaaaataagg 300
 taccatgctc tttatta 317

<210> 454

<211> 231

<212> DNA

<213> Homo sapiens

<400> 454

ttcgaggtag aatcaactct cagagtgtag tttccttcta tagatgagtc agcattaata 60
 taagccacgc cacgctcttg aaggagtctt gaattctcct ctgctcactc agtagaacca 120
 agaagaccaa attcttctgc atcccagctt gcaaacaaaa ttgttcttct aggtctccac 180
 cttcctttt tcagtgttcc aaagctcctc acaatttcat gaacaacagc t 231

<210> 455

<211> 231

<212> DNA

<213> Homo sapiens

<400> 455

taccaaagag ggcataataa tcagtctcac agtaggggtc accatcctcc aagtgaaaaa 60
 cattgttccg aatgggcttt ccacaggcta cacacacaaa acaggaaaca tgccaagtgtt 120
 gtttcaacgc attgatgact tctccaagga tcttcctttg gcatcgacca cattcagggg 180
 caaagaattht ctcatagcac agtcacaaat acagggtctc tttctcctct a 231

<210> 456
 <211> 231
 <212> DNA
 <213> Homo sapiens

<400> 456
 ttggcaggta cccttacaaa gaagacacca taccttatgc gttattaggt ggaataatca 60
 ttccattcag tattatcggt attattcttg gagaaaccct gtctgtttac tgtaaccttt 120
 tgcactcaaa ttccctttatc aggaataaact acatagccac tatttacaaa gccattggaa 180
 cctttttatt tgggtgcagct gctagtcagt ccctgactga cattgccaag t 231

<210> 457
 <211> 231
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(231)
 <223> n = A,T,C or G

<400> 457
 cgaggtagcc aggggtctga aaatctctnn ttantagtc gatagcaaaa ttgttcatca 60
 gcattcctta atatgatctt gctataatta gatttttctc cattagagtt catacagttt 120
 tatttgattt tattagcaat ctctttcaga agacccttga gatcattaag ctttgtatcc 180
 agttgtctaa atcgatgcct catttcctct gaggtgtcgc tggcttttgt g 231

<210> 458
 <211> 231
 <212> DNA
 <213> Homo sapiens

<400> 458
 aggtctggtt cccccactt ccactccct ctactctctc taggactggg ctgggcccaag 60
 agaagagggg tgggttaggga agccgttgag acctgaagcc ccacctcta ccttccttca 120
 acaccctaac cttgggtaac agcatttgga attatcattt gggatgagta gaatttccaa 180
 ggtcctgggt taggcatttt gggggggccag accccaggag aagaagattc t 231

<210> 459
 <211> 231
 <212> DNA
 <213> Homo sapiens

<400> 459
 ggtaccgagg ctgcgtgaca cagagaaacc ccaacgcgag gaaaggaatg gccagccaca 60
 ctttcgcgaa acctgtggtg gccaccagt cctaacggga caggacagag agacagagca 120
 gccctgact gttttccctc caccacagcc atcctgtccc tcattggctc tgtgctttcc 180
 actatacaca gtcaccgtcc caatgagaaa caagaaggag caccctccac a 231

<210> 460
 <211> 231

<212> DNA

<213> Homo sapiens

<400> 460

```
gcaggtataa catgctgcaa caacagatgt gactaggaac ggccggtgac atggggaggg 60
cctatcaccc tattcttggg ggctgcttct tcacagtgat catgaagcct agcagcaaat 120
cccacctccc cacacgcaca cggccagcct ggagcccaca gaagggtcct cctgcagcca 180
gtggagcttg gtccagcctc cagtccaccc ctaccaggct taaggataga a 231
```

<210> 461

<211> 231

<212> DNA

<213> Homo sapiens

<400> 461

```
cgaggtttga gaagctctaa tgtgcagggg agccgagaag caggcggcct agggaggggtc 60
gcgtgtgctc cagaagagtg tgtgcatgcc agaggggaaa caggcgcttg tgtgtcctgg 120
gtgggggttca gtgaggagtg ggaaattggg tcagcagAAC caagccgttg ggtgaataag 180
agggggattc catggcactg atagagccct atagtttcag agctgggaat t 231
```

<210> 462

<211> 231

<212> DNA

<213> Homo sapiens

<400> 462

```
aggtaccctc attgtagcca tgggaaaatt gatgttcagt ggggatcagt gaattaaatg 60
gggtcatgca agtataaaaa ttaaaaaaaa aagacttcat gcccaatctc atatgatgtg 120
gaagaactgt tagagagacc aacagggtag tgggttagag atttcagag tcttacattt 180
tctagaggag gtatttaatt tcttctcact catccagtgt tgtatttagg a 231
```

<210> 463

<211> 231

<212> DNA

<213> Homo sapiens

<400> 463

```
tactccagcc tggtgacaga gcgagaccct atcacgccc cccacccac caaaaaaaaa 60
actgagtaga caggtgtcct cttggcatgg taagtcttaa gtcccctccc agatctgtga 120
catttgacag gtgtcttttc ctctggacct cgggtgtccc atctgagtga gaaaaggcag 180
tggggaggtg gatcttccag tcgaagcggg atagaagccc gtgtgaaaag c 231
```

<210> 464

<211> 231

<212> DNA

<213> Homo sapiens

<400> 464

```
gtactctaag attttatcta agttgccttt tctgggtggg aaagttaaac cttagtgact 60
aaggacatca catatgaaga atgtttaagt tggaggtggc aacgtgaatt gcaaacaggg 120
cctgcttcag tgactgtgtg cctgtagtcc cagctactcg ggagtctgtg tgaggccagg 180
```


ggtgccagcg caccagctag atgctctgta acttctagggc cccattttcc c 231

<210> 465

<211> 231

<212> DNA

<213> Homo sapiens

<400> 465

catgttgttg tagctgtggt aatgctggct gcatctcaga caggggtaac ttcagctcct 60
gtggcaaatt agcaacaaat tctgacatca tatttatggg ttctgtatct ttgttgatga 120
aggatggcac aattttttgct tgtgttcata atatactcag attagttcag ctccatcaga 180
taaactggag acatgcagga cattagggta gtgttgtagc tctggtaatg a 231

<210> 466

<211> 231

<212> DNA

<213> Homo sapiens

<400> 466

caggtacctc tttccattgg atactgtgct agcaagcatg ctctccgggg tttttttaat 60
ggccttcgaa cagaacttgc cacataccca ggtataatag tttctaacat ttgccagga 120
cctgtgcaat caaatattgt ggagaattcc ctagctggag aagtcacaaa gactataggc 180
aataatggag accagtccca caagatgaca accagtcgtt gtgtgcggct g 231

<210> 467

<211> 311

<212> DNA

<213> Homo sapiens

<400> 467

gtacaccctg gcacagtcca atctgaactg gttcggcact catctttcat gagatggatg 60
tggtggcttt tctccttttt catcaagact cctcagcagg gagcccagac cagcctgcac 120
tgtgccttaa cagaaggctc tgagattcta agtgggaatc atttcagtga ctgtcatgtg 180
gcatgggtct ctgccaagc tctgaatgag actatagcaa ggcggctgtg ggacgtcagt 240
tgtgacctgc tgggcctccc aatagactaa caggcagtgc cagttggacc caagagaaga 300
ctgcagcaga c 311

<210> 468

<211> 3112

<212> DNA

<213> Homo sapiens

<400> 468

cattgtgttg ggagaaaaac agagggggaga tttgtgtggc tgcagccgag ggagaccagg 60
aagatctgca tgggtgggaag gacctgatga tacagagttt gataggagac aattaaaggc 120
tggaagycac tggatgcctg atgatgaagt ggactttcaa actggggcac tactgaaacg 180
atgggatggc cagagacaca ggagatgagt tggagcaagc tcaataacaa agtgggttcaa 240
cgaggacttg gaattgcatg gagctggagc tgaagtttag cccaattggt tactagttaga 300
gtgaatgtgg atgattggat gatcatttct catctctgag cctcagggtc cccatccata 360
aaatgggata cacagtatga tctataaagt gggatatagt atgatctact tcactggggtt 420
atttgaagga tgaattgaga taattttattt cagggtgccta gaacaatgcc cagattagta 480

```

catttggtgg aactgagaaa tggcataaca ccaaatttaa tatatgtcag atgttactat 540
gattatcatt caatctcata gttttgtcat ggcccaattt atcctcactt gtgcctcaac 600
aaattgaact gttaacaaaag gaatctctgg tectgggtaa tggctgagca cactgagca 660
tttccattcc agttggcttc ttgggtttgc tagctgcac actagtcac ttaaataaat 720
gaagttttaa catttctcca gtgatttttt tatctcacct ttgaagatac tatgttatgt 780
gattaaataa agaacttgag aagaacaggt ttcattaaac ataaaatcaa tgtagacgca 840
aattttctgg atgggcaata cttatgttca caggaaatgc tttaaaatat gcagaagata 900
attaaatggc aatggacaaa gtgaaaaact tagacttttt tttttttttt ggaagtatct 960
ggatgttctt tagtcactta aaggagaact gaaaaatagc agtgagtctc acataatcca 1020
acctgtgaga ttaaggctct ttgtggggaa ggacaaagat ctgtaaattt acagtttctt 1080
tccaaagcca acgtcgaatt ttgaaacata tcaaagctct tcttcaagac aaataatcta 1140
tagtacatct ttcttatggg atgcacttat gaaaaatggg ggctgtcaac atctagtcac 1200
tttagctctc aaaatggttc attttaagag aaagttttag aatctcata tttattctgt 1260
ggaaggacag cattgtggct tggactttat aaggctctta ttcaactaaa taggtgagaa 1320
ataagaaagg ctgctgactt taccatctga ggccacacat ctgctgaaat ggagataatt 1380
aacatcacta gaaacagcaa gatgacaata taatgtctaa gtagtacat gtttttgcac 1440
atttccagcc cctttaaata tccacacaca cagggaagc aaaaggaagc acagagatcc 1500
ctgggagaaa tgcccgccg ccatctggg tcatogatga gctcgcct gtgcctggct 1560
ccgcttgta gggaaggaca ttagaaaatg aattgatgtg ttccttaaag gatgggcagg 1620
aaaacagatc ctgttggtga tatttatgtg aacgggatta cagatttgaa atgaagtcac 1680
aaagtgagca ttaccaatga gaggaaaaca gacgagaaaa tcttgatggc ttcacaagac 1740
atgcaacaaa caaaatggaa tactgtgatg acatgaggca gccaaagctg ggaggagata 1800
accacggggc agagggtcag gattctggcc ctgctgccta aactgtgcgt tcataaccaa 1860
atcatttcat atttctaacc ctcaaaaaca agctgttgta atatctgatc tctacggttc 1920
cttctgggcc caacattctc catatatcca gccacactca tttttaatat ttagttccca 1980
gatctgtact gtgacctttc tacactgtag aataacatta ctcatthtgt tcaaagacc 2040
ttcgtgttgc tgcctaatat gtagctgact gtttttcta aggagtgtt tggcccagg 2100
gatctgtgaa caggctggga agcatctcaa gatctttcca gggttatact tactagcaca 2160
cagcatgatc attacggagt gaattatcta atcaacatca tctcagtggt ctttgcceat 2220
actgaaattc atttcccact tttgtgcca ttctcaagac ctcaaaatgt cattccatta 2280
atatcacagg attaaacttt ttttttaacc tggaagaatt caatgttaca tgcagctatg 2340
ggaatttaac tacataattt gttttccagt gcaaagatga ctaagtcctt tatccctccc 2400
ctttgtttga ttttttttcc agtataaagt taaaatgctt agccttgtag tgaggctgta 2460
tacagccaca gcctctcccc atccctccag ccttatctgt catcaccatc aaccctccc 2520
atgcacctaa acaaaatcta acttgtaatt ccttgaacat gtcaggcata cattattcct 2580
tctgcctgag aagctcttcc ttgtctctta aatctagaat gatgtaaagt tttgaataag 2640
ttgactatct tacttcatgc aaagaaggga cacatatgag attcatcatc acatgagaca 2700
gcaaatacta aaagtgtaat ttgattataa gagtttagat aaatatatga aatgcaagag 2760
ccacagaggg aatgtttatg gggcacgttt gtaagcctgg gatgtgaagc aaaggcagg 2820
aacctcatag tatcttatat aatatacttc atttctctat ctctatcaca atatccaaca 2880
agcttttcac agaattcatg cagtgc aaat ccccaaagg aacctttatc catttcatgg 2940
tgagtgcgct ttagaatttt ggcaaatcat actggtcact tatctcaact ttgagatgtg 3000
tttgtccttg tagttaattg aaagaaatag ggcactcttg tgagccactt taggggtcac 3060
tcttggaat aaagaattta caaagagcaa aaaaaaaaaa aaaaaaaaaa aa 3112

```

<210> 469

<211> 2229

<212> DNA

<213> Homo sapiens

<400> 469

agctctttgt aaattcttta ttgccaggag tgaaccctaa agtgggtcac aagagtgtccc 60
 tatttctttc aattaactac aaggacaaac acatctcaaa gttgagataa gtgaccagta 120
 tgatttgcca aaattctaaa gcgcactcac catgaaatgg ataaagggtta cctttgggga 180
 tttgcactgc atgaattctg tgaaaagcct gttggatatt gtgatagaga tagagaaatg 240
 aagtatatta tataagatac tatgagggtc cctgcctttg cttcacatcc caggcttaca 300
 aacgtgtccc ataaacattc cctctgtggc tcttgcatct catatattta tctaaactct 360
 tataatcaaa tacactttta gtatttgctg tctcatgtga tgatgaatct catatgtgtc 420
 ccttctttgc atgaagtaag atagtcaact tattcaaaac tttacatcat tctagattta 480
 agagacaagg aagagcttct caggcagaag gaataatgta tgcctgacat gttcaaggaa 540
 ttacaagtta gattttgttt aggtgcatgg gaggggttga tgggtgatgac agataaggct 600
 ggaggggatg ggagaggctg tggctgtata cagcctcagt acaaggctaa gcattttaac 660
 tttatactgg aaaaaaaatc aaacaaaggg gagggataaa ggacttagtc atctttgcac 720
 tggaaaacaa aatatgtaat taaattccca tagctgcatg taacattgaa ttcttccagg 780
 ttaaaaaaaa agttaatcct gtgatattaa tggaaatgaca ttttgaggtc ttgagaatgg 840
 gcacaaaagt gggaaatgaa tttcagtatg ggcaaagaca ctgaggatga tgttgattag 900
 ataattcact ccgtaatgat catgctgtgt gctagtaagt ataaccctgg aaagatcttg 960
 agatgcttcc cagcctgttc acagatcccc tgggccagaa cactccttag gaaaaacagt 1020
 cagctacata ttaggcagca acacgaaggg tctttgaaca aaatgagtaa tgttattcta 1080
 cagtgtagaa aggtcacagt acagatctgg gaactaaata ttaaaaatga gtgtggctgg 1140
 atatatggag aatgttgggc ccagaaggaa ccgtagagat cagatattac aacagctttg 1200
 ttttgagggt tagaaatatg aaatgatttg gttatgaacg cacagtttag gcagcagggc 1260
 cagaatcctg accctctgcc ccgtgggtat ctctcccca gcttggctgc ctcatgtcat 1320
 cacagtattc cattttgttt gttgcatgtc ttgtgaagcc atcaagattt tctcgtctgt 1380
 tttcctctca ttgtaatgc tctctttgtg acttcatttc aaatctgtaa tcccgttcaa 1440
 ataaatatcc acaacaggat ctgttttctt gccatcctt taaggaacac atcaattcat 1500
 tttctaattg ccttccctca caagcgggac caggcacagg gcgaggctca tcgatgaccc 1560
 aagatggcgg ccgggcattt ctcccaggga tctctgtgct tccttttgtg cttcctgtgt 1620
 gtgtggatat ttaaaggggc tggaaatgtg caaaaacatg tctactacta gacattatat 1680
 tgtcatcttg ctgtttctag tgatgttaat tatctccatt tcagcagatg tgtggcctca 1740
 gatggtaaag tcagcagcct ttcttatttc tcacctggaa atacatacga ccatttgagg 1800
 agacaaatgg caaggtgtca gcataccctg aacttgagtt gagagctaca cacaatatta 1860
 ttggtttccg agcatcacia acaccctctc tgtttcttca ctgggcacag aattttaata 1920
 cttatttcag tgggctgttg gcaggaacaa atgaagcaat ctacataaag tcaactagtgc 1980
 agtgccctgac acacaccatt ctcttgaggc cccctctaga gatcccacag gtcatatgac 2040
 ttcttgggga gcagtggctc acacctgtaa tcccagcact ttgggaggct gaggcagggt 2100
 ggtcacctga ggtcaggagt tcaagaccag cctggccaat atggtgaaac cccatctcta 2160
 ctaaaaatc aaaaattagc tgggcgtgct ggtgcatgcc tgtaatccca gcccacac 2220
 aatggaatt 2229

<210> 470

<211> 2426

<212> DNA

<213> Homo sapiens

<400> 470

gtaaatctt tattgccagg agtgaaccct aaagtggctc acaagagtgc cctatttctt 60
 tcaattaaact acaaggacaa acacatctca aagttgagat aagtgaccag tatgatttgc 120
 caaaattcta aagcgactc accatgaaat ggataaagggt tacctttggg gatttgcact 180
 gcatgaattc tgtgaaaagc ttgttgata ttgtgataga gatagagaaa tgaagtatat 240
 tatataagat actatgagggt tccctgcctt tgcttcacat cccaggctta caaacgtgcc 300
 ccataaacat tccctctgtg gctcttgcac ttcatatatt tatctaaact cttataatca 360

```

aattacactt ttagtatttg ctgtctcatg tgatgatgaa tctcatatgt gtcccttctt 420
tgcatagaat aagatagtc aattattcaa aactttacat cattctagat ttaagagaca 480
aggaagagct tctcaggcag aaggaataat gtatgcctga catgttcaag gaattacaag 540
ttagattttg tttaggtgca tgggaggggt tgatggtgat gacagataag gctggaggga 600
tggggagagg ctgtggctgt atacagcctc agtacaaggc taagcatttt aactttatac 660
tggaaaaaaa atcaaacaaa ggggagggat aaaggactta gtcactcttg cactggaaaa 720
caaaatatgt aattaaattc ccatagctgc atgtaacatt gaattcttcc aggttaaaaa 780
aaaaagttaa tcctgtgata ttaatggaat gacattttga ggtcttgaga atgggcacaa 840
aagtgggaaa tgaatttcag tatgggcaa gacactgagg atgatgttga ttagataatt 900
cactccgtaa tgatcatgct gtgtgctagt aagtataacc ctggaaagat cttgagatgc 960
ttcccagcct gttcacagat cccctgggcc agaacactcc ttaggaaaaa cagtcagcta 1020
catattaggc agcaacacga aggtctttg aacaaaatga gtaatgttat tctacagtgt 1080
agaaaggcca cagtcagat ctgggaacta aatattaaaa atgagtgtgg ctggatata 1140
ggagaatggt gggcccagaa ggaaccgtag agatcagata ttacaacagc tttgttttga 1200
gggttagaaa tatgaaatga tttggttatg aacgcacagt ttaggcagca gggccagaat 1260
cctgaccctc tgccccgtgg ttatctctc cccagcttgg ctgcctcatg tcatcacagt 1320
attccatttt gtttgttgca tgtcttgta agccatcaag atttctcgt ctgttttctt 1380
ctcattggta atgctcactt tgtgacttca tttcaaactc gtaatccgt tcaaataaat 1440
atccacaaca ggatctgttt tcctgcccac cctttaagga acacatcaat tcatcttcta 1500
atgtccttcc ctcacaagcg ggaccaggca cagggcgagg ctcacgatg acccaagatg 1560
gcggccgggc atttctcca gggatctctg tgettcttt tgtgcttct gtgtgtgtgg 1620
atatttaaag gggctggaaa tgtgcaaaaa catgtcacta cttagacatt atattgtcat 1680
cttgcgtgtt ctagtgatgt taattatctc catttcagca gatgtgtggc ctcagatgg 1740
aaagtcagca gcctttctta tttctcacct ggaaatacat acgaccattt gaggagacaa 1800
atggcaaggt gtcagcatac cctgaacttg agttgagagc tacacacaat attattggtt 1860
tccgagcatc acaaacaccc tctctgtttc ttcactgggc acagaatttt aatacttatt 1920
tcagtgggct gttggcagga acaaatgaag caatctacat aaagtcacta gtgcagtgcc 1980
tgacacacac cattctcttg aggtccctc tagagatccc acaggtcata tgacttcttg 2040
gggagcagtg gctcacacct gtaatcccag cactttggga ggctgaggca ggtgggtcac 2100
ctgaggtcag gagttcaaga ccagcctggc caatatggtg aaaccccatc tctactaaaa 2160
atacaaaaat tagctgggag tgcctgtgca tgctgtaat cccagctact tgggaggctg 2220
aggcaggaga attgctggaa catgggaggc ggaggttgca gtgagctgta attgtgccat 2280
tgactcga cctgggcgac agagtggaa tctgtttcca aaaaacaaac aaacaaaaaa 2340
ggcatagtc gataaacgt ggggtgggat tgtaaataga agcaggatat aaagggcag 2400
gggtgacggt tttgccaac acaatg

```

<210> 471

<211> 812

<212> DNA

<213> Homo sapiens

<400> 471

```

gaacaaaatg agtaatgtta ttctacagtg tagaaaggct acagtacaga tctgggaact 60
aaatattaaa aatgagtggt gctggatata tggagaatgt tgggccaga aggaaccgta 120
gagatcagat attacaacag ctttgttttg agggttagaa atatgaaatg atttggttat 180
gaacgcacag tttaggcagc agggccagaa tcctgaccct ctgccccgtg gttatctcct 240
ccccagcttg gctgcctcat gtcacacag tattccattt tgtttgttgc atgtcttg 300
aagccatcaa gatcttctcg tctgttttcc tctcattggt aatgctcact ttgtgacttc 360
atttcaaatt tgtaatcccc ttcaaataaa tatccacaac aggatctgtt ttctgcccc 420
tcctttaagg aacacatcaa ttcattttct aatgtccttc cctcacaagc gggaccaggc 480
acagggcgag gctcatcgat gacccaagat ggcggccggg catttctccc agggatctct 540

```

```

gtgcttcctt ttgtgcttcc tgtgtgtgtg gatattttaa ggggctggaa atgtgcaaaa 600
acatgtcact acttagacat tatattgtca tcttgctgtt tctagtgatg ttaattatct 660
ccatttcagc agatgtgtgg cctcagatgg taaagtcagc agcctttctt atttctcacc 720
tctgtatcat caggtccttc ccaccatgca gatcttctcg gtctccctcg gctgcagcca 780
cacaaatctc ccctctgttt ttctgatgcc ag 812

```

<210> 472

<211> 515

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1) ... (515)

<223> n = A,T,C or G

<400> 472

```

acggagactt attttctgat attgtctgca tatgtatgtt ttttaagagtc tggaaatagt 60
cttatgactt tcctatcatg cttattaata aataatacag cccagagaag atgaaaatgg 120
gttcagaat tatttgctct tgcagcccg tgaatctcag caagaggaac caccaactga 180
caatcaggat attgaacctg gacaagagag agaaggaaca cctccgatcg aagaacgtaa 240
agtagaaggt gattgccagg aaatggatct ggaaaagact cggagtgagc gtggagatgg 300
ctctgatgta aaagagaaga ctccacctaa tctaagcat gctaagacta aagaagcagg 360
agatgggcag ccataagtta aaaagaagac aagctgaagc tacacacatg gctgatgtca 420
cattgaaaat gtgactgaaa atttgaaaat tctctcaata aagtttgagt tttctctgaa 480
gaaaaaaaaa naaaaaaaaa aaanaaaan aaaaa 515

```

<210> 473

<211> 5829

<212> DNA

<213> Homo sapiens

<400> 473

```

cgcattgccg ggaagcccaa gctggctcga agagccacca gccacctgtg caaggggtggg 60
cctggaccag ttggaccagc caccaagctc acctactcaa ggaagcaggg atggccagg 120
tgcaacagcc tgagtggctg ccacctgata gctgatggag cagaggcctg aggaaaatca 180
gatggcacat ttagctcttt aatggatctt aagttaattt ttctataaag cacatggcac 240
cagtcacatg ctcagagctc gtatggcact ggcgaccaca gcaggccgag ttcccaggat 300
tgccatccag gggggccttc tgtagccctg gccagacctt gcagaggtgg ctgggtgctc 360
tttgagcgag ctgcgcctcc ctggcatgca caggccccag gtactgacac gctgctctga 420
gtgagcttgt cctgccttgg ctgccacctt actgctgatg gagcagcggc cttaggaaaa 480
gcaaatggcg ctgtagccca actttagggt agaagaagat gtaccatgtc cggccgctag 540
ttggtgactg gtgcacctgc tcttgccgta cccttgccaga ggtgggtggg tgctctttgg 600
ccagcttggc ctgacctggc atgcacaagc ctcaagtcaa caactgtcct acaaatggag 660
acacagagag gaaacaagca ggcggctcag gagcaggggtg tgtgctgcct ttggggctcc 720
agtccatgcc tcgggtcgta tgggtactga ggcttcttgg ttgccaagag gcggaccaca 780
ggccttcttg aggaggactt tacgttcaag tgcagaaagc agccaaaatt accatccatg 840
agactaagcc ttctgtggcc ctggcgagac ttaaaatttg tgccaaggca ggacaagctc 900
actcggagca gcgtgtcagt agctggggcc tatgcatgcc gggcagggcc gggctggctg 960
aaggagcaac cagccacctc tgcaaggggt cgcctagtgc aggcggagca tccaccacct 1020
caccgcctcg aggaagtggg gatggccagg ttcccacagc ctgagtgtct gccaccttat 1080

```



```

actgtcacat ggacacacca tcacacacat gaacacaccg acacactgcc atatggacac 960
tggcacacac actgccacac tgtcacatgg acacacctcc acaccatcac accaccacac 1020
acactgcctg tggacacaag gacacacaga cactgtcaca cagatacaca aaacactgtc 1080
acacgggagac atcaccatgc agatacacca ccactctggg gccgtctgaa ttaccctgct 1140
gggggggacag cagtggcata ctcatgccta agtgactggc tttcacccca gtagtgattg 1200
ccctccatca acactgcca cccaggttg gggctacccc agcccatctt tacaaaacag 1260
ggcaagggtga actaatggag tgggtggagg agttggaaga aatcccagcg tcagtcaccg 1320
ggatagaatt cccaaggaac cctctttttg gaggatgggt tccatttctg gaggcgatct 1380
gccgacaggg tgaatgcctt cttgcttgct ttctggggaa tcagagagag tccgttttgt 1440
ggtgggaaga gtgtggctgt gtactttgaa ctctgtaaa ttctctgact catgtccaca 1500
aaaccaacag ttttgtgaat gtgtctggag gcaagggaag ggccactcag gatctatggt 1560
gaagggaaga ggctggggc tggagtattc gctt 1594

```

<210> 475

<211> 2414

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (33)

<223> n=A,T,C or G

<400> 475

```

cccaacacaa tggctttata agaatgcttc acntgtgaaa aacaaatata aaagtcttct 60
tgtagattat ttttaaggac aaatctttat tccatgttta atttatttag ctttccctgt 120
agctaataat tcatgctgaa cacattttaa atgctgtaaa tgtagataat gtaatttatg 180
tatcattaat gcctcttttag tagtttagag aaaacgtcaa aagaaatggc ccagaataa 240
gcttcttgat ttgtaaaatt ctatgtcatt ggctcaaatt tgtatagtat ctcaaaatat 300
aaatatatag acatctcaga taatatattt gaaatagcaa attcctgtta gaaaataata 360
gtacttaact agatgagaat aacaggctgc cattatttga attgtctcct attcgttttt 420
catttgttgt gttactcatg ttttacttat ggggggatat atataacttc cgctgttttc 480
agaagtattg tatgcagtca gtatgagaat gcaatttaag tttccttgat gctttttcac 540
acttctatta ctagaaataa gaatacagta atattggcaa agaaaattga ccagttcaat 600
aaaatttttt agtaaactctg attgaaaata aacattgctt atggctttct tacatcaata 660
ttgttatgtc ctagacacct tatctgaaat tacggcttca aaattctaata tatgtgcaaa 720
tgtgtaaaat atcaatactt tatgttcaag ctggggcctc ttcaggcgctc ctgggctgag 780
agagaaagat gctagctccg caagccgggg agggaacacc gccacattgt tacatggaca 840
caccgccacg tggacacatg accagactca catgtacaga cacacggaga cattaccaca 900
tgagacacac gtcacacagt cacacgagca cactggcata gtcacatgga cggacacaca 960
gacatatgga gaaatcacac tgacacacca ccacactate acagggacac agacacacgg 1020
agacatcacc acatggacac actgtcacac taccacaggg acacgagaca tcacactgtc 1080
acatggacac accatcacac acatgaacac accgacacac tgccatatgg acactgccac 1140
acacactgcc acactgtcac atggacacac ctccatacca tcacaccacc acacacactg 1200
ccatgtggac acaaggacac acagacactg tcacacagat acacaaaaca ctgtcacacg 1260
gagacatcac catgcagata caccaccaca tggacatagc accagacact ctgccacaca 1320
gatacaccac cacacagaaa tgccggacaca ctgccacaca gacaccacca catcgttgcc 1380
acactttcat gtgtcagctg gcggtgtggg cccacagact ctgggctcta atcgagaaat 1440
tacttggaca tatagtgaag gcaaaatttt tttttatttt ctgggtaacc aagcgcgact 1500
ctgtctcaaa aaaagaaaaa aaaagcaata tactgtgtaa tcgttgacag cataattcac 1560
tattatgtag atcggagagc agaggattct gaatgcatga acatatcatt aacatttcaa 1620

```



```

tacattactc ataattactg atgaactaaa gagaaaccaa gaaattatgg tgatagttat 1680
attgacctgg agaaatgtag acacaaaaga accgtaagat gagaaatgtg ttaacacagt 1740
ctataagggc atgcaagaat aaaaataggg gagaaaacag gagagttttt caagagcttt 1800
ctggcatgt aagtcaactt gtatcgggta atttttaaaa ggttttattt catgcaataa 1860
actgcacata cttcaattgt acatttttgg aattcttggc atttgtagct ctataaaacc 1920
agcaacatat taaaatagca aacatatcca ttacctttac caccaaagtt ttcttgtgtt 1980
ttttctactc actttttcct gcctatcccc ccctctcttc cacaggtaac cactgatcca 2040
cttccagtca ctatccatga gtttttattt ccaaatacat gaaatcatat gaatttctgg 2100
tttttctgt tggagcccaa ggagcaaggg cagaatgagg aacatgatgt ttcttwccga 2160
cagttactca tgacgtctcc atccaggact gaggggggca tccttctcca tctaggactg 2220
ggggcatcct tctccatcca gtattggggg tcctccttct ccattccagta ttgggggtca 2280
tcctcctcca tccaggacct gaggggtgtc cttttctgog cttccttgga tggcagctctt 2340
tccttcatg tttatagtra cttaccatta aatcactgtg ccgttttttc ctaaaataaa 2400
aaaaaaaaaa aaaa 2414

```

<210> 476

<211> 3434

<212> DNA

<213> Homo sapiens

<400> 476

```

ctgtgctgca aatggggcca tatagaggaa aggagcagct ggctctggag catggtgtgc 60
actccctttg ggcttcagtt ccattgtctc tgggtcgtat gacactgcgg gcttgttggg 120
tgccaagagg cagaccacag gtcattctga ggaggacttt atgttccagt ccagaaagca 180
gccagtggta ccaccagggt gacttgtgct tctgtggccc aggccagacg tagaatttga 240
caaagtcagg acggtctcag tcagagcagc atgtcgggtc ccggggcctg tgcattgccg 300
gcagggccag gctggcttaa ggagcaagca gccacctctg ttagggtgtg gcctggagca 360
ggtggagcag ccaccaacct cacgcactga aagaagcagg gatggccagg ttccaacatc 420
ctgagtggct gccacctgat ggctgatgga gcagaggcct gagggaaagc agatggcact 480
gctttgtagt gctgttcttt gtctctcttg atctttttca gttaatgtct gttttatcag 540
agactaggat tgcaaaccct gctctttttt gctttccatt tgcttggtaa atattcctcc 600
atccctttat tttaaacctt tgtgtgtctt tgcacatgag atgggtctcc tgaatacagg 660
acaacaatgg gtctttactc tttatccaac ttgccagttc gtgtctttta actggggcat 720
ttagccatt tacatttaag ttttagtattt gttacatgtg aaatttatcc tgtcatgatg 780
ttgctagctt tttatttttc ccattagttt gcagtttctt tatagtgtca atggctctta 840
caattcgata tgtttttgta gtggctggta ctgggttttc ctttctacgt ttagtgtctc 900
cttcaggagc tcttgtaaca caagaatgtg gatatttttc ttgtaaggta aatatgtgga 960
tttattctgg gactgtattc tatggccttt accccaagaa tcattacttt ttaaaatgca 1020
attcaaatta gcataaaaca tttacagcct atggaaaggc ttgtggcatt agaatcctta 1080
tttataggat tatttttgtt ttttttgaga tatggtcttt gtcattcagg cagaagtgcc 1140
gtggtttgat cataattcac cacagccctg aactcttgag tccaagccat ctttttgcct 1200
taatctccca accagttgga tctacaagca taaggcatca tgcgtggcta attttttcac 1260
gttttttttt tttttgtcga gattatggta tcaactgtgt gctctggctg atctcaaagt 1320
tttgacctca agggatcttt ctgccacagc ctctaaagt gctaggatta tatgcatgat 1380
acaccatgcc tattgtagag tattacatta ttttcaaagt cttattgtaa gagcatttta 1440
ttgccttttg cctaaataac tcaatataat atctctgaaa cttttttttg acaaattttg 1500
gggcgtgatg atgagagaag ggggtttgaa acttttcta aagagttaac ttagagccat 1560
ttaagaaagg aaaaaacaca aattatcaga aaaacaacag taagatcaag tgcaaaagtt 1620
ctgtggcaaa gatgatgaga gtaaagaata tatgtttgtg actcatgggt gcttttactt 1680
tgttcttgaa tttctgagta cgggttaaca tttaaagaat ctacattata gataacattt 1740
tattgcaagt aaatgtattt caaaatttgt tattggtttt gtatgagatt attctcagcc 1800

```

```

<210> 477
<211> 141
<212> PRT
<213> Homo sapiens

<400> 477
Met Asp Gly His Thr Asp Ile Trp Arg Asn His Met Asp Thr Pro Pro
          5                      10                      15

His Tyr His Arg Asp Thr Asp Thr Arg Arg His His His Met Asp Thr
          20                      25                      30

Leu Ser His Tyr His Arg Asp Thr Arg His His Thr Val Thr Trp Thr
          35                      40                      45

His His His Thr His Glu His Thr Asp Thr Leu Pro Tyr Gly His Trp
          50                      55                      60

His Thr His Cys His Thr Val Thr Trp Thr His Leu His Thr Ile Thr
          65                      70                      75                      80

Pro Pro His Thr Leu Pro Val Asp Thr Arg Thr His Arg His Cys His

```

				85					90					95	
Thr	Asp	Thr	Gln	Asn	Thr	Val	Thr	Arg	Arg	His	His	His	Ala	Asp	Thr
			100					105					110		
Pro	Pro	Leu	Trp	Cys	Arg	Leu	Asn	Tyr	Pro	Ala	Gly	Gly	Thr	Ala	Val
		115					120					125			
Ala	Tyr	Ser	Cys	Leu	Ser	Asp	Trp	Leu	Ser	Pro	Gln				
	130					135					140				

<210> 478

<211> 144

<212> PRT

<213> Homo sapiens

<400> 478

Met	Tyr	Arg	His	Thr	Glu	Thr	Leu	Pro	His	Gly	Asp	Thr	Val	Thr	Gln
				5					10					15	
Ser	His	Gly	His	Thr	Gly	Ile	Val	Thr	Trp	Thr	Asp	Thr	Gln	Thr	Tyr
			20					25					30		
Gly	Glu	Ile	Thr	Trp	Thr	His	His	His	Thr	Ile	Thr	Gly	Thr	Gln	Thr
		35				40						45			
His	Gly	Asp	Ile	Thr	Thr	Trp	Thr	His	Cys	His	Thr	Thr	Thr	Gly	Thr
	50					55					60				
Arg	Asp	Ile	Thr	Leu	Ser	His	Gly	His	Thr	Ile	Thr	His	Met	Asn	Thr
	65				70					75				80	
Pro	Thr	His	Cys	His	Met	Asp	Thr	Gly	Thr	His	Thr	Ala	Thr	Leu	Ser
				85					90					95	
His	Gly	His	Thr	Ser	Thr	Pro	Ser	His	His	Thr	His	Cys	Leu	Trp	
		100						105				110			
Thr	Gln	Gly	His	Thr	Asp	Thr	Val	Thr	Gln	Ile	His	Lys	Thr	Leu	Ser
		115					120					125			
His	Gly	Asp	Ile	Thr	Met	Gln	Ile	His	His	His	Ser	Gly	Ala	Val	
	130					135					140				

<210> 479

<211> 223

<212> PRT

<213> Homo sapiens

<400> 479

Met Tyr Arg His Thr Glu Thr Leu Pro His Gly Asp Thr Val Thr Gln
 5 10 15

Ser His Glu His Thr Gly Ile Val Thr Trp Thr Asp Thr Gln Thr Tyr
 20 25 30

Gly Glu Ile Thr Leu Thr His His His Thr Ile Thr Gly Thr Gln Thr
 35 40 45

His Gly Asp Ile Thr Thr Trp Thr His Cys His Thr Thr Thr Gly Thr
 50 55 60

Arg Asp Ile Thr Leu Ser His Gly His Thr Ile Thr His Met Asn Thr
 65 70 75 80

Pro Thr His Cys His Met Asp Thr Ala Thr His Thr Ala Thr Leu Ser
 85 90 95

His Gly His Thr Ser Ile Pro Ser His His His Thr His Cys His Val
 100 105 110

Asp Thr Arg Thr His Arg His Cys His Thr Asp Thr Gln Asn Thr Val
 115 120 125

Thr Arg Arg His His His Ala Asp Thr Pro Pro His Gly His Ser Thr
 130 135 140

Arg His Ser Ala Thr Gln Ile His His His Thr Glu Met Arg Thr His
 145 150 155 160

Cys His Thr Asp Thr Thr Thr Ser Leu Pro His Phe His Val Ser Ala
 165 170 175

Gly Gly Val Gly Pro Thr Thr Leu Gly Ser Asn Arg Glu Ile Thr Trp
 180 185 190

Thr Tyr Ser Glu Gly Lys Ile Phe Phe Tyr Phe Leu Gly Asn Gln Ala
 195 200 205

Arg Leu Cys Leu Lys Lys Arg Lys Lys Lys Gln Tyr Thr Val
 210 215 220

<210> 480

<211> 145

<212> PRT

<213> Homo sapiens

<400> 480

Met Glu Pro Tyr Arg Gly Asn Glu Gln Pro Ser Gln Glu Gln Gly Val

5 10 15
 Cys Cys Leu Trp Gly Leu Gln Ser Leu Pro Gln Gly Ser Tyr Val Thr
 20 25 30
 Val Gly Phe Leu Val Val Lys Arg Gln Thr Ile Gly Arg Leu Glu Arg
 35 40 45
 Asp Phe Met Phe Lys Cys Arg Lys Gln Pro Gly Leu Pro Pro Ser Gly
 50 55 60
 Leu Cys Leu Leu Trp Pro Trp Pro Asn Leu Glu Phe Gly Arg Arg Gln
 65 70 75 80
 Asp Arg Leu Thr Trp Ser Ser Val Ser Val Ala Gly Val Cys Ala Cys
 85 90 95
 Arg Ala Arg Pro Gly Trp Leu Gly Glu Gln Pro Ala Thr Ser Ala Gly
 100 105 110
 Val Arg Leu Glu Gln Val Glu Gln Pro Pro Ala His Pro Leu Gln Glu
 115 120 125
 Ala Gly Val Ala Arg Phe Pro Arg Pro Glu Trp Val Pro Pro Asn Gly
 130 135 140

<210> 481

<211> 168

<212> PRT

<213> Homo sapiens

<400> 481

Met His Gly Pro Gln Val Leu Ala Arg Cys Ser Glu Cys Ala Cys Pro
 5 10 15
 Ala Leu Ala Ala Thr Ser Ala Gly Val Arg Leu Glu Gly Val Asp Arg
 20 25 30
 Pro Pro Thr Leu Pro Ser Gln Gly Ser Gly Trp Pro Cys Ser His Ser
 35 40 45
 Leu Ser Gly Cys His Leu Met Ala Asp Gly Ala Lys Ala Leu Gly Lys
 50 55 60
 Ala Asp Gly Pro Trp Pro Tyr Leu Phe Val Arg Arg Thr Asp Val Pro
 65 70 75 80

<210> 483
 <211> 144
 <212> PRT
 <213> Homo sapiens

<400> 483
 Met Glu Thr Gln Arg Gly Asn Lys Gln Arg Ala Gln Glu Gln Gly Val
 5 10 15
 Cys Cys Leu Trp Gly Ser Ser Pro Cys Leu Gly Ser Tyr Gly Thr Ala
 20 25 30
 Gly Phe Leu Val Ala Lys Arg Arg Thr Thr Gly Leu Leu Glu Glu Asp
 35 40 45
 Phe Thr Phe Lys Cys Arg Lys Gln Pro Lys Leu Pro Ser Met Arg Leu
 50 55 60
 Ser Leu Leu Trp Pro Trp Arg Asp Leu Lys Phe Val Pro Arg Gln Asp
 65 70 75 80
 Lys Leu Thr Arg Ser Ser Val Ser Val Ala Gly Ala Tyr Ala Cys Arg
 85 90 95
 Ala Gly Pro Gly Trp Leu Lys Glu Gln Pro Ala Thr Ser Ala Arg Val
 100 105 110
 Arg Leu Val Gln Ala Glu His Pro Pro Pro His Pro Leu Glu Glu Val
 115 120 125
 Gly Met Ala Arg Phe Pro Gln Pro Glu Cys Leu Pro Pro Tyr Cys
 130 135 140

<210> 484
 <211> 30
 <212> PRT
 <213> Homo Sapien

<400> 484
 Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe
 1 5 10 15
 Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile
 20 25 30

<210> 485
 <211> 31
 <212> DNA
 <213> Artificial Sequence

<220>

<223> Made in a lab

<400> 485

gggaagctta tcacctatgt gccgcctctg c

31

<210> 486

<211> 27

<212> DNA

<213> Artificial Sequence

<220>

<223> Made in a lab

<400> 486

gcgaattctc acgctgagta tttggcc

27

<210> 487

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> Made in a lab

<400> 487

cccgaattct tagctgcca tccgaacgcc ttcac

36

<210> 488

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Made in a lab

<400> 488

gggaagcttc ttccccggct gcaccagctg tgc

33

<210> 489

<211> 19

<212> PRT

<213> Artificial Sequence

<220>

<223> Made in a lab

<400> 489

Met Asp Arg Leu Val Gln Arg Phe Gly Thr Arg Ala Val Tyr Leu Ala

1

5

10

15

Ser Val Ala

<210> 490
 <211> 20
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 490
 Tyr Leu Ala Ser Val Ala Ala Phe Pro Val Ala Ala Gly Ala Thr Cys
 1 5 10 15
 Leu Ser His Ser
 20

<210> 491
 <211> 20
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 491
 Thr Cys Leu Ser His Ser Val Ala Val Val Thr Ala Ser Ala Ala Leu
 1 5 10 15
 Thr Gly Phe Thr
 20

<210> 492
 <211> 20
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 492
 Ala Leu Thr Gly Phe Thr Phe Ser Ala Leu Gln Ile Leu Pro Tyr Thr
 1 5 10 15
 Leu Ala Ser Leu
 20

<210> 493
 <211> 20
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<211> 20
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 497
 Leu Leu Pro Pro Pro Pro Ala Leu Cys Gly Ala Ser Ala Cys Asp Val
 1 5 10 15
 Ser Val Arg Val
 20

<210> 498
 <211> 20
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 498
 Asp Val Ser Val Arg Val Val Val Gly Glu Pro Thr Glu Ala Arg Val
 1 5 10 15
 Val Pro Gly Arg
 20

<210> 499
 <211> 20
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 499
 Arg Val Val Pro Gly Arg Gly Ile Cys Leu Asp Leu Ala Ile Leu Asp
 1 5 10 15
 Ser Ala Phe Leu
 20

<210> 500
 <211> 20
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 500
 Leu Asp Ser Ala Phe Leu Leu Ser Gln Val Ala Pro Ser Leu Phe Met

1 5 10 15
 Gly Ser Ile Val
 20
 <210> 501
 <211> 20
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> Made in a lab
 <400> 501
 Phe Met Gly Ser Ile Val Gln Leu Ser Gln Ser Val Thr Ala Tyr Met
 1 5 10 15
 Val Ser Ala Ala
 20

<210> 502
 <211> 414
 <212> DNA
 <213> Homo Sapien

<400> 502
 caccatggag acaggcctgc gctggctttt cctggctcgt gtgctcaaag gtgtccaatg 60
 tcagtccgtg gaggagtccg ggggtcgct ggtcacgcct gggacacctt tgacantcac 120
 ctgtagagtt tttggaatng acctcagtag caatgcaatg agctgggtcc gccaggctcc 180
 agggaagggg ctggaatgga tcggagccat tgataattgt ccacantacg cgacctgggc 240
 gaaaggccga ttnatnatntt ccaaaacctn gaccacggtg gatttgaaaa tgaccagtcc 300
 gacaaccgag gacacggcca cctatntttg tggcagaatg aatactggta atagtgggtg 360
 gaagaatatt tggggcccag gcaccctggt caccgtntcc tcagggcaac ctaa 414

<210> 503
 <211> 379
 <212> DNA
 <213> Homo Sapien

<400> 503
 atncgatggt gcttggtcaa aggtgtccag tgtcagtcgg tggaggagtc cgggggtcgc 60
 ctggtcacgc ctgggacacc cctgacactc acctgcaccg tntctggatt ngacatcagt 120
 agctatggag tgagctgggt ccgccaggct ccagggaagg ggctggnata catcgatca 180
 ttagtagtag tggtagatntt tacgcgagc gggcgaaagg ccgattcacc atttccaaaa 240
 cctngaccac ggtggatttg aaaatcacca gtttgacaac cgaggacacg gccacctatt 300
 tntgtgccag agggggggtt aattataaag acatttgggg ccagggcacc ctggtcaccg 360
 tntccttagg gcaacctaa 379

<210> 504
 <211> 19
 <212> PRT
 <213> Artificial Sequence


```

1              5              10              15

    <210> 512
    <211> 15
    <212> PRT
    <213> Artificial Sequence

    <220>
    <223> Made in a lab

    <400> 512
Asp Ser Gly Gly Pro Leu Ile Cys Asn Gly Tyr Leu Gln Gly Leu
1              5              10              15

    <210> 513
    <211> 15
    <212> PRT
    <213> Artificial Sequence

    <220>
    <223> Made in a lab

    <400> 513
Ala Pro Cys Gly Gln Val Gly Val Pro Asx Val Tyr Thr Asn Leu
1              5              10              15

    <210> 514
    <211> 15
    <212> PRT
    <213> Artificial Sequence

    <220>
    <223> Made in a lab

    <400> 514
Leu Cys Lys Phe Thr Glu Trp Ile Glu Lys Thr Val Gln Ala Ser
1              5              10              15

    <210> 515
    <211> 15
    <212> PRT
    <213> Artificial Sequence

    <220>
    <223> Made in a lab

    <400> 515
Met Val Glu Ala Ser Leu Ser Val Arg His Pro Glu Tyr Asn Arg
1              5              10              15

    <210> 516

```

<211> 15
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 516
 Val Ser Glu Ser Asp Thr Ile Arg Ser Ile Ser Ile Ala Ser Gln
 1 5 10 15

<210> 517
 <211> 15
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 517
 Glu Val Cys Ser Lys Leu Tyr Asp Pro Leu Tyr His Pro Ser Met
 1 5 10 15

<210> 518
 <211> 15
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 518
 Arg Ala Glu Pro Gly Thr Glu Ala Arg Arg His Tyr Asp Glu Gly
 1 5 10 15

<210> 519
 <211> 17
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 519
 Arg Ala Glu Pro Gly Thr Glu Ala Arg Arg Asn Tyr Asp Glu Gly Cys
 1 5 10 15
 Gly

<210> 520
 <211> 25

<212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 520
 Val Gly Glu Gly Leu Tyr Gln Gly Val Pro Arg Ala Glu Pro Gly Thr
 1 5 10 15
 Glu Ala Arg Arg His Tyr Asp Glu Gly
 20 25

<210> 521
 <211> 21
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 521
 Ala Pro Phe Pro Asn Gly His Val Gly Ala Gly Gly Ser Gly Leu Leu
 1 5 10 15
 Pro Pro Pro Pro Ala
 20

<210> 522
 <211> 20
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 522
 Leu Leu Val Val Pro Ala Ile Lys Lys Asp Tyr Gly Ser Gln Glu Asp
 1 5 10 15
 Phe Thr Gln Val
 20

<210> 523
 <211> 254
 <212> PRT
 <213> Artificial Sequence

<220>
 <223> Made in a lab

<400> 523
 Met Ala Thr Ala Gly Asn Pro Trp Gly Trp Phe Leu Gly Tyr Leu Ile
 1 5 10 15

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Leu Gly Val Ala Gly Ser Leu Val Ser Gly Ser Cys Ser Gln Ile Ile
 20 25 30
 Asn Gly Glu Asp Cys Ser Pro His Ser Gln Pro Trp Gln Ala Ala Leu
 35 40 45
 Val Met Glu Asn Glu Leu Phe Cys Ser Gly Val Leu Val His Pro Gln
 50 55 60
 Trp Val Leu Ser Ala Thr His Cys Phe Gln Asn Ser Tyr Thr Ile Gly
 65 70 75 80
 Leu Gly Leu His Ser Leu Glu Ala Asp Gln Glu Pro Gly Ser Gln Met
 85 90 95
 Val Glu Ala Ser Leu Ser Val Arg His Pro Glu Tyr Asn Arg Pro Leu
 100 105 110
 Leu Ala Asn Asp Leu Met Leu Ile Lys Leu Asp Glu Ser Val Ser Glu
 115 120 125
 Ser Asp Thr Ile Arg Ser Ile Ser Ile Ala Ser Gln Cys Pro Thr Ala
 130 135 140
 Gly Asn Ser Cys Leu Val Ser Gly Trp Gly Leu Leu Ala Asn Gly Arg
 145 150 155 160
 Met Pro Thr Val Leu Gln Cys Val Asn Val Ser Val Val Ser Glu Glu
 165 170 175
 Val Cys Ser Lys Leu Tyr Asp Pro Leu Tyr His Pro Ser Met Phe Cys
 180 185 190
 Ala Gly Gly Gly Gln Xaa Gln Xaa Asp Ser Cys Asn Gly Asp Ser Gly
 195 200 205
 Gly Pro Leu Ile Cys Asn Gly Tyr Leu Gln Gly Leu Val Ser Phe Gly
 210 215 220
 Lys Ala Pro Cys Gly Gln Val Gly Val Pro Gly Val Tyr Thr Asn Leu
 225 230 235 240
 Cys Lys Phe Thr Glu Trp Ile Glu Lys Thr Val Gln Ala Ser
 245 250

<210> 524

<211> 765

<212> DNA

<213> Homo sapien

<400> 524

```

atggccacag caggaaatcc ctggggctgg ttcttgggggt acctcatcct tgggtgtcgca      60
ggatcgctcg tctctggtag ctgcagccaa atcataaacg gcgaggactg cagcccgcac      120
tcgcagccct ggcaggcggc actggtcatg gaaaacgaat tggttctgctc gggcgtcctg      180
gtgcatccgc agtgggtgct gtcagccgca cactgtttcc agaactccta caccatcggy      240
ctgggcctgc acagtcttga ggccgaccaa gagccaggga gccagatggt ggaggccagc      300
ctctccgtac ggcacccaga gtacaacaga cccttgctcg ctaacgacct catgctcatc      360
aagttagacg aatccgtgtc cgagtctgac accatccgga gcatcagcat tgcttcgcag      420
tgccctaccg cggggaactc ttgcctcggt tctggctggg gtctgctggc gaacggcaga      480
atgcctaccg tgctgcagtg cgtgaacgtg tcggtggtgt ctgaggaggt ctgcagtaag      540
ctctatgacc cgctgtacca cccagcatg ttctgcgccg gcggagggca agaccagaag      600
gactcctgca acggtgactc tggggggccc ctgatctgca acgggtactt gcagggcctt      660
gtgtctttcg gaaaagcccc gtgtggccaa gttggcgtgc caggtgtcta caccaacctc      720
tgcaaatcca ctgagtggat agagaaaacc gtccaggcca gttaa              765

```

<210> 525
 <211> 254
 <212> PRT
 <213> Homo sapien

<400> 525
 Met Ala Thr Ala Gly Asn Pro Trp Gly Trp Phe Leu Gly Tyr Leu Ile
 1 5 10 15
 Leu Gly Val Ala Gly Ser Leu Val Ser Gly Ser Cys Ser Gln Ile Ile
 20 25 30
 Asn Gly Glu Asp Cys Ser Pro His Ser Gln Pro Trp Gln Ala Ala Leu
 35 40 45
 Val Met Glu Asn Glu Leu Phe Cys Ser Gly Val Leu Val His Pro Gln
 50 55 60
 Trp Val Leu Ser Ala Ala His Cys Phe Gln Asn Ser Tyr Thr Ile Gly
 65 70 75 80
 Leu Gly Leu His Ser Leu Glu Ala Asp Gln Glu Pro Gly Ser Gln Met
 85 90 95
 Val Glu Ala Ser Leu Ser Val Arg His Pro Glu Tyr Asn Arg Pro Leu
 100 105 110
 Leu Ala Asn Asp Leu Met Leu Ile Lys Leu Asp Glu Ser Val Ser Glu
 115 120 125
 Ser Asp Thr Ile Arg Ser Ile Ser Ile Ala Ser Gln Cys Pro Thr Ala
 130 135 140
 Gly Asn Ser Cys Leu Val Ser Gly Trp Gly Leu Leu Ala Asn Gly Arg
 145 150 155 160
 Met Pro Thr Val Leu Gln Cys Val Asn Val Ser Val Val Ser Glu Glu
 165 170 175
 Val Cys Ser Lys Leu Tyr Asp Pro Leu Tyr His Pro Ser Met Phe Cys
 180 185 190
 Ala Gly Gly Gly Gln Asp Gln Lys Asp Ser Cys Asn Gly Asp Ser Gly
 195 200 205
 Gly Pro Leu Ile Cys Asn Gly Tyr Leu Gln Gly Leu Val Ser Phe Gly
 210 215 220
 Lys Ala Pro Cys Gly Gln Val Gly Val Pro Gly Val Tyr Thr Asn Leu
 225 230 235 240
 Cys Lys Phe Thr Glu Trp Ile Glu Lys Thr Val Gln Ala Ser
 245 250

<210> 526
 <211> 963
 <212> DNA
 <213> Homo sapiens

<400> 526
 atgagttcct gcaacttcac acatgccacc tttgtgctta ttggtatccc aggattagag 60
 aaagcccatt tctgggttgg cttccccctc ctttccatgt atgtagtggc aatgttttga 120
 aactgcatcg tgggtcttcat cgtaaggacg gaacgcagcc tgcacgctcc gatgtacctc 180
 tttctctgca tgcttgcagc cattgacctg gccttatcca catccaccat gcctaagatc 240
 cttgcccttt tctgggtttga ttcccgagag attagctttg aggcctgtct taccagatg 300
 ttctttattc atgccctctc agccattgaa tccaccatcc tgctggccat ggcctttgac 360

```

cgttatgtgg ccactctgcc cccactgcgc catgctgcag tgetcaacaa tacagtaaca 420
gcccagattg gcatcgtagc tgtggtccgc ggatccctct tttttttccc actgcctctg 480
ctgatcaagc ggctggcctt ctgccactcc aatgtcctct cgcactccta ttgtgtccac 540
caggatgtaa tgaagttagc ctatgcagac actttgcccc atgtggtata tgggtcttact 600
gccattctgc tggatcatgg cgtggacgta atgttcatct ccttgtccta ttttctgata 660
atacgaacgg ttctgcaact gccttccaag tcagagcggg ccaaggcctt tggaacctgt 720
gtgtcacaca ttggtgtggt actcgccttc tatgtgccac ttattggcct ctcagttgta 780
caccgctttg gaaacagcct tcattcccatt gtgcgtgttg tcattgggtga catctacctg 840
ctgctgcctc ctgtcatcaa tcccatcatc tatggtgcc aacacaaaca gatcagaaca 900
cgggtgctgg ctatgttcaa gatcagctgt gacaaggact tgcaggctgt gggaggcaag 960
tga

```

<210> 527

<211> 321

<212> PRT

<213> Homo sapiens

<400> 527

```

Met Ser Ser Cys Asn Phe Thr His Ala Thr Phe Val Leu Ile Gly Ile
          5                      10                      15

```

```

Pro Gly Leu Glu Lys Ala His Phe Trp Val Gly Phe Pro Leu Leu Ser
          20                      25                      30

```

```

Met Tyr Val Val Ala Met Phe Gly Asn Cys Ile Val Val Phe Ile Val
          35                      40                      45

```

```

Arg Thr Glu Arg Ser Leu His Ala Pro Met Tyr Leu Phe Leu Cys Met
          50                      55                      60

```

```

Leu Ala Ala Ile Asp Leu Ala Leu Ser Thr Ser Thr Met Pro Lys Ile
          65                      70                      75                      80

```

```

Leu Ala Leu Phe Trp Phe Asp Ser Arg Glu Ile Ser Phe Glu Ala Cys
          85                      90                      95

```

```

Leu Thr Gln Met Phe Phe Ile His Ala Leu Ser Ala Ile Glu Ser Thr
          100                     105                     110

```

```

Ile Leu Leu Ala Met Ala Phe Asp Arg Tyr Val Ala Ile Cys His Pro
          115                     120                     125

```

```

Leu Arg His Ala Ala Val Leu Asn Asn Thr Val Thr Ala Gln Ile Gly
          130                     135                     140

```

```

Ile Val Ala Val Val Arg Gly Ser Leu Phe Phe Phe Pro Leu Pro Leu
          145                     150                     155                     160

```

```

Leu Ile Lys Arg Leu Ala Phe Cys His Ser Asn Val Leu Ser His Ser
          165                     170                     175

```

Tyr Cys Val His Gln Asp Val Met Lys Leu Ala Tyr Ala Asp Thr Leu
180 185 190

Pro Asn Val Val Tyr Gly Leu Thr Ala Ile Leu Leu Val Met Gly Val
195 200 205

Asp Val Met Phe Ile Ser Leu Ser Tyr Phe Leu Ile Ile Arg Thr Val
210 215 220

Leu Gln Leu Pro Ser Lys Ser Glu Arg Ala Lys Ala Phe Gly Thr Cys
225 230 235 240

Val Ser His Ile Gly Val Val Leu Ala Phe Tyr Val Pro Leu Ile Gly
245 250 255

Leu Ser Val Val His Arg Phe Gly Asn Ser Leu His Pro Ile Val Arg
260 265 270

Val Val Met Gly Asp Ile Tyr Leu Leu Leu Pro Pro Val Ile Asn Pro
275 280 285

Ile Ile Tyr Gly Ala Lys Thr Lys Gln Ile Arg Thr Arg Val Leu Ala
290 295 300

Met Phe Lys Ile Ser Cys Asp Lys Asp Leu Gln Ala Val Gly Gly Lys
305 310 315 320

<210> 528

<211> 20

<212> DNA

<213> Homo Sapien

<400> 528

actatggtcc agaggctgtg

20

<210> 529

<211> 20

<212> DNA

<213> Homo Sapien

<400> 529

atcacctatg tgccgcctct

20

<210> 530

<211> 1852

<212> DNA

<213> Homo sapiens

<400> 530

ggcacgagaa ttaaaacct cagcaaaaca ggcatagaag ggacatacct taaagtaata 60

aaaaccacct atgacaagcc cacagccaac ataatactaa atggggaaaa gttagaagca 120

```

tttctctga gaactgcaac aataaataca aggatgctgg attttgtcaa atgccttttc 180
tgtgtctgtt gagatgctta tgtgactttg cttttaattc tgtttatgtg attatcacat 240
ttattgactt gcctgtgtta gaccggaaga gctggggtgt ttctcaggag ccaccgtgtg 300
ctgcggcagc ttccgggataa cttgaggctg catcactggg gaagaaacac aytccctgtcc 360
gtggcgctga tggctgagga cagagcttca gtgtggcttc tctgogactg gcttcttcgg 420
ggagtcttc cttcatagtt catccatatg gctccagagg aaaattatat tattttgtta 480
tggatgaaga gtattacgtt gtgcagatat actgcagtgt cttcatctct tgatgtgtga 540
ttgggtaggt tccaccatgt tgccgcagat gacatgattt cagtacctgt gtctggctga 600
aaagtgtttg tttgtgaatg gatattgtgg tttctggatc tcatcctctg tgggtggaca 660
gctttctcca ccttgctgga agtgacctgc tgtccagaag tttgatggct gaggagtata 720
ccatcgtgca tgcacttttc atttcctgca tttcttctc cctggatgga cagggggagc 780
ggcaagagca acgtgggcac ttctggagac cacaacgact cctctgtgaa gacgcttggg 840
agcaagaggt gcaagtgggt ctgccactgc tccccctgct gcagggggag cggcaagagc 900
aacgtggtcg cttggggaga ctacgatgac agcgcttca tggatcccag gtaccacgtc 960
catggagaag atctggacaa gctccacaga gctgcctggg ggggtaaagt ccccagaaag 1020
gatctcatcg tcatgctcag ggacacggat gtgaacaaga gggacaagca aaagaggact 1080
gctctacatc tggcctctgc caatgggaat tcagaagtag taaaactcgt gctggacaga 1140
cgatgtcaac ttaatgtcct tgacaacaaa aagaggacag ctctgacaaa ggccgtacaa 1200
tgccaggaag atgaatgtgc gttaatgttg ctggaacatg gcactgatcc aaatattcca 1260
gatgagtatg gaaataccac tctacactat gctgtctaca atgaagataa attaatggcc 1320
aaagcactgc tcttatacgg tgctgatata gaatcaaaaa acaagcatgg cctcacacca 1380
ctgctacttg gtatacatga gcaaaaacag caagtgggtga aatttttaat caagaaaaaa 1440
gcgaatttaa atgcgctgga tagatatgga agaactgctc tcatacttgc tgtatgttgt 1500
ggatcagcaa gtatagtcag cctctacttt gagcaaaatg ttgatgtatc ttctcaagat 1560
ctggaaagac ggccagagag tatgctgttt ctagtcatca tcatgtaatt tgccagttac 1620
tttctgacta caaagaaaaa cagatgttaa aaatctcttc tgaaaacagc aatccagaac 1680
aagacttaaa gctgacatca gaggaagagt cacaaaggct taaaggaagt gaaaacagcc 1740
agccagagct agaagattta tggctattga agaagaatga agaacacgga agtactcatg 1800
tgggattccc agaaaacctg actaacggtg ccgctgctgg caatggtgat ga 1852

```

<210> 531

<211> 879

<212> DNA

<213> Homo sapiens

<400> 531

```

atgcactctt catttctctgc atttcttctc ccttggatgg acagggggag cggcaagagc 60
aacgtgggca cttctggaga ccacaacgac tctctgtga agacgcttgg gagcaagagg 120
tgcaagtggg gctgccactg cttccccctg tgcaggggga gcggcaagag caacgtgggtc 180
gcttggggag actacgatga cagcgcttcc atggatccca ggtaccacgt ccatggagaa 240
gatctggaca agctccacag agctgcctgg tggggtaaag tccccagaaa ggatctcatc 300
gtcatgctca gggacacgga tgtgaacaag agggacaagc aaaagaggac tgctctacat 360
ctggcctctg ccaatgggaa ttcagaagta gtaaaactcg tgctggacag acgatgtcaa 420
cttaatgtcc ttgacaacaa aaagaggaca gctctgacaa aggccgtaca atgccaggaa 480
gatgaatgtg cgtaaatgtt gctggaacat ggcactgatc caaatattcc agatgagtat 540
ggaaatacca ctctacacta tgctgtctac aatgaagata aattaatggc caaagcactg 600
ctcttatacg gtgctgatat cgaatcaaaa aacaagcatg gcctcacacc actgctactt 660
ggtatacatg agcaaaaaca gcaagtgggt aaatttttaa tcaagaaaaa agcgaattta 720
aatgcgctgg atagatatgg aagaactgct ctcatacttg ctgtatgttg tggatcagca 780
agtatagtca gccctctact tgagcaaaat gttgatgtat cttctcaaga tctggaaaaga 840
cggccagaga gtatgctgtt tctagtcatc atcatgtaa 879

```


Asn Ala Leu Asp Arg Tyr Gly Arg Thr Ala Leu Ile Leu Ala Val Cys
 245 250 255

Cys Gly Ser Ala Ser Ile Val Ser Pro Leu Leu Glu Gln Asn Val Asp
 260 265 270

Val Ser Ser Gln Asp Leu Glu Arg Arg Pro Glu Ser Met Leu Phe Leu
 275 280 285

Val Ile Ile Met
 290

<210> 533

<211> 801

<212> DNA

<213> Homo sapiens

<400> 533

```
atgtacaagc ttccagtgcac caactgtgct acaaattggag ccacagagag gaaacaagca 60
gcaggctcag gagcagggtta tgcgctgcct tcggctctcc aatccatgcc tcagggtccc 120
tatgccactg cagcattctt ggttgccaag aggccaacca caggccatct tgagaaggag 180
tttatgttcc actgcagaaa gcagccagga tcaccatcca ggggacttgg tcttctgtgg 240
ccctggccag acatagaatt tgtgccaagg caggacaagc tctctcagag cagcgtgtta 300
gtacctcaaa tctgtgcgtg ccagacaagg ccaaactggc tcaatgagca accagccacc 360
tctgcagggg tgcgtctgga ggaggtggac cagccaccaa ccttaccag tcaaggaagt 420
ggatggccat gttccacacag cctgagtggc tgccacctga tggctgatat agcaaaggcc 480
ttaggaaaag cagatggccc ttggccctac ctttttgta gaagaaccga tgttccatgt 540
cctgcagcga gtgaggttg tggctgtgcc ccagctcct ggcacaccct cgcagagggtg 600
actggttgct ctttgagccc tcttagcctt gccagcatg cacaagcctc agtgctacta 660
ctgtgctaca aatggagcca tataggggaa acgagcagcc atctcaggag caaggtgtat 720
gctgccttgg ggggctccag tccttgcttc aagggtctta tgtcactgtg ggcttcttgg 780
ttgccaagag gcagaccata g                                     801
```

<210> 534

<211> 267

<212> PRT

<213> Homo sapiens

<400> 534

Met Tyr Lys Leu Gln Cys Asn Asn Cys Ala Thr Asn Gly Ala Thr Glu
 5 10 15

Arg Lys Gln Ala Ala Gly Ser Gly Ala Gly Tyr Ala Leu Pro Ser Ala
 20 25 30

Leu Gln Ser Met Pro Gln Gly Ser Tyr Ala Thr Ala Arg Phe Leu Val
 35 40 45

Ala Lys Arg Pro Thr Thr Gly His Leu Glu Lys Glu Phe Met Phe His

50 55 60
 Cys Arg Lys Gln Pro Gly Ser Pro Ser Arg Gly Leu Gly Leu Leu Trp
 65 70 75 80
 Pro Trp Pro Asp Ile Glu Phe Val Pro Arg Gln Asp Lys Leu Thr Gln
 85 90 95
 Ser Ser Val Leu Val Pro Gln Ile Cys Ala Cys Gln Thr Arg Pro Asn
 100 105 110
 Trp Leu Asn Glu Gln Pro Ala Thr Ser Ala Gly Val Arg Leu Glu Glu
 115 120 125
 Val Asp Gln Pro Pro Thr Leu Pro Ser Gln Gly Ser Gly Trp Pro Cys
 130 135 140
 Ser His Ser Leu Ser Gly Cys His Leu Met Ala Asp Ile Ala Lys Ala
 145 150 155 160
 Leu Gly Lys Ala Asp Gly Pro Trp Pro Tyr Leu Phe Val Arg Arg Thr
 165 170 175
 Asp Val Pro Cys Pro Ala Ala Ser Glu Val Gly Gly Cys Ala Pro Ser
 180 185 190
 Ser Trp His Thr Leu Ala Glu Val Thr Gly Cys Ser Leu Ser Pro Leu
 195 200 205
 Ser Leu Ala Gln His Ala Gln Ala Ser Val Leu Leu Leu Cys Tyr Lys
 210 215 220
 Trp Ser His Ile Gly Glu Thr Ser Ser His Leu Arg Ser Lys Val Tyr
 225 230 235 240
 Ala Ala Phe Gly Gly Ser Ser Pro Cys Leu Lys Gly Leu Met Ser Leu
 245 250 255
 Trp Ala Ser Trp Leu Pro Arg Gly Arg Pro
 260 265

<210> 535
 <211> 6082
 <212> DNA
 <213> Homo sapiens

<400> 535
 cctccactat tacagcttat aggaaattac aatccacttt acaggcctca aaggttcatt 60
 ctggcccgagc ggacaggcgt ggcggccgga gccccagcat ccctgcttga ggtccaggag 120
 cggagccccgc ggccactgcc gcctgatcag cgcgaccccc gcccgcgcc gccccgccg 180

180	185	190
Gln Ile Val Asn Leu Leu Ser	Asn Asp Val Asn Lys Phe Asp Gln Val	
195	200	205
Thr Val Phe Leu His Phe Leu Trp	Ala Gly Pro Leu Gln Ala Ile Ala	
210	215	220
Val Thr Ala Leu Leu Trp Met Glu Ile Gly	Ile Ser Cys Leu Ala Gly	
225	230	235
Met Ala Val Leu Ile Ile Leu Leu Pro	Leu Gln Ser Cys Phe Gly Lys	
	245	250
Leu Phe Ser Ser Leu Arg Ser Lys Thr	Ala Thr Phe Thr Asp Ala Arg	
	260	265
Ile Arg Thr Met Asn Glu Val Ile Thr Gly	Ile Arg Ile Ile Lys Met	
	275	280
Tyr Ala Trp Glu Lys Ser Phe Ser Asn Leu	Ile Thr Asn Leu Arg Lys	
	290	295
Lys Glu Ile Ser Lys Ile Leu Arg Ser Ser	Cys Leu Arg Gly Met Asn	
305	310	315
Leu Ala Ser Phe Phe Ser Ala Ser Lys Ile	Ile Val Phe Val Thr Phe	
	325	330
Thr Thr Tyr Val Leu Leu Gly Ser Val Ile	Thr Ala Ser Arg Val Phe	
	340	345
Val Ala Val Thr Leu Tyr Gly Ala Val Arg	Leu Thr Val Thr Leu Phe	
	355	360
Phe Pro Ser Ala Ile Glu Arg Val Ser Glu	Ala Ile Val Ser Ile Arg	
	370	375
Arg Ile Gln Thr Phe Leu Leu Leu Asp Glu	Ile Ser Gln Arg Asn Arg	
385	390	395
Gln Leu Pro Ser Asp Gly Lys Lys Met Val	His Val Gln Asp Phe Thr	
	405	410
Ala Phe Trp Asp Lys Ala Ser Glu Thr Pro	Thr Leu Gln Gly Leu Ser	
	420	425
Phe Thr Val Arg Pro Gly Glu Leu Leu Ala	Val Val Gly Pro Val Gly	
	435	440
Ala Gly Lys Ser Ser Leu Leu Ser Ala Val	Leu Gly Glu Leu Ala Pro	

450		455		460
Ser His Gly Leu Val	Ser Val His Gly Arg	Ile Ala Tyr Val	Ser Gln	
465	470	475	480	
Gln Pro Trp Val	Phe Ser Gly Thr Leu Arg	Ser Asn Ile Leu	Phe Gly	
	485	490	495	
Lys Lys Tyr Glu Lys	Glu Arg Tyr Glu Lys	Val Ile Lys Ala	Cys Ala	
	500	505	510	
Leu Lys Lys Asp Leu	Gln Leu Leu Glu Asp	Gly Asp Leu Thr	Val Ile	
	515	520	525	
Gly Asp Arg Gly Thr	Thr Leu Ser Gly Gly	Gln Lys Ala Arg	Val Asn	
	530	535	540	
Leu Ala Arg Ala Val	Tyr Gln Asp Ala Asp	Ile Tyr Leu Leu	Asp Asp	
545	550	555	560	
Pro Leu Ser Ala Val	Asp Ala Glu Val Ser	Arg His Leu Phe	Glu Leu	
	565	570	575	
Cys Ile Cys Gln Ile	Leu His Glu Lys Ile	Thr Ile Leu Val	Thr His	
	580	585	590	
Gln Leu Gln Tyr Leu	Lys Ala Ala Ser Gln	Ile Leu Ile Leu	Lys Asp	
	595	600	605	
Gly Lys Met Val Gln	Lys Gly Thr Tyr Thr	Glu Phe Leu Lys	Ser Gly	
	610	615	620	
Ile Asp Phe Gly Ser	Leu Leu Lys Lys Asp	Asn Glu Glu Ser	Glu Gln	
625	630	635	640	
Pro Pro Val Pro Gly	Thr Pro Thr Leu Arg	Asn Arg Thr Phe	Ser Glu	
	645	650	655	
Ser Ser Val Trp Ser	Gln Gln Ser Ser Arg	Pro Ser Leu Lys	Asp Gly	
	660	665	670	
Ala Leu Glu Ser Gln	Asp Thr Glu Asn Val	Pro Val Thr Leu	Ser Glu	
	675	680	685	
Glu Asn Arg Ser Glu	Gly Lys Val Gly Phe	Gln Ala Tyr Lys	Asn Tyr	
	690	695	700	
Phe Arg Ala Gly Ala	His Trp Ile Val Phe	Ile Phe Leu Ile	Leu Leu	
705	710	715	720	
Asn Thr Ala Ala Gln	Val Ala Tyr Val Leu	Gln Asp Trp Trp	Leu Ser	

725										730				735			
Tyr	Trp	Ala	Asn	Lys	Gln	Ser	Met	Leu	Asn	Val	Thr	Val	Asn	Gly	Gly		
740				745				750									
Gly	Asn	Val	Thr	Glu	Lys	Leu	Asp	Leu	Asn	Trp	Tyr	Leu	Gly	Ile	Tyr		
755				760				765									
Ser	Gly	Leu	Thr	Val	Ala	Thr	Val	Leu	Phe	Gly	Ile	Ala	Arg	Ser	Leu		
770				775				780									
Leu	Val	Phe	Tyr	Val	Leu	Val	Asn	Ser	Ser	Gln	Thr	Leu	His	Asn	Lys		
785				790				795				800					
Met	Phe	Glu	Ser	Ile	Leu	Lys	Ala	Pro	Val	Leu	Phe	Phe	Asp	Arg	Asn		
805				810				815									
Pro	Ile	Gly	Arg	Ile	Leu	Asn	Arg	Phe	Ser	Lys	Asp	Ile	Gly	His	Leu		
820				825				830									
Asp	Asp	Leu	Leu	Pro	Leu	Thr	Phe	Leu	Asp	Phe	Ile	Gln	Thr	Leu	Leu		
835				840				845									
Gln	Val	Val	Gly	Val	Val	Ser	Val	Ala	Val	Ala	Val	Ile	Pro	Trp	Ile		
850				855				860									
Ala	Ile	Pro	Leu	Val	Pro	Leu	Gly	Ile	Ile	Phe	Ile	Phe	Leu	Arg	Arg		
865				870				875				880					
Tyr	Phe	Leu	Glu	Thr	Ser	Arg	Asp	Val	Lys	Arg	Leu	Glu	Ser	Thr	Thr		
885				890				895									
Arg	Ser	Pro	Val	Phe	Ser	His	Leu	Ser	Ser	Ser	Leu	Gln	Gly	Leu	Trp		
900				905				910									
Thr	Ile	Arg	Ala	Tyr	Lys	Ala	Glu	Glu	Arg	Cys	Gln	Glu	Leu	Phe	Asp		
915				920				925									
Ala	His	Gln	Asp	Leu	His	Ser	Glu	Ala	Trp	Phe	Leu	Phe	Leu	Thr	Thr		
930				935				940									
Ser	Arg	Trp	Phe	Ala	Val	Arg	Leu	Asp	Ala	Ile	Cys	Ala	Met	Phe	Val		
945				950				955				960					
Ile	Ile	Val	Ala	Phe	Gly	Ser	Leu	Ile	Leu	Ala	Lys	Thr	Leu	Asp	Ala		
965				970				975									
Gly	Gln	Val	Gly	Leu	Ala	Leu	Ser	Tyr	Ala	Leu	Thr	Leu	Met	Gly	Met		
980				985				990									
Phe	Gln	Trp	Cys	Val	Arg	Gln	Ser	Ala	Glu	Val	Glu	Asn	Met	Met	Ile		

```

          995              1000              1005
Ser Val Glu Arg Val Ile Glu Tyr Thr Asp Leu Glu Lys Glu Ala Pro
  1010              1015              1020

Trp Glu Tyr Gln Lys Arg Pro Pro Pro Ala Trp Pro His Glu Gly Val
  1025              1030              1035              1040

Ile Ile Phe Asp Asn Val Asn Phe Met Tyr Ser Pro Gly Gly Pro Leu
          1045              1050              1055

Val Leu Lys His Leu Thr Ala Leu Ile Lys Ser Gln Glu Lys Val Gly
          1060              1065              1070

Ile Val Gly Arg Thr Gly Ala Gly Lys Ser Ser Leu Ile Ser Ala Leu
  1075              1080              1085

Phe Arg Leu Ser Glu Pro Glu Gly Lys Ile Trp Ile Asp Lys Ile Leu
  1090              1095              1100

Thr Thr Glu Ile Gly Leu His Asp Leu Arg Lys Lys Met Ser Ile Ile
  1105              1110              1115              1120

Pro Gln Glu Pro Val Leu Phe Thr Gly Thr Met Arg Lys Asn Leu Asp
          1125              1130              1135

Pro Phe Asn Glu His Thr Asp Glu Glu Leu Trp Asn Ala Leu Gln Glu
          1140              1145              1150

Val Gln Leu Lys Glu Thr Ile Glu Asp Leu Pro Gly Lys Met Asp Thr
          1155              1160              1165

Glu Leu Ala Glu Ser Gly Ser Asn Phe Ser Val Gly Gln Arg Gln Leu
  1170              1175              1180

Val Cys Leu Ala Arg Ala Ile Leu Arg Lys Asn Gln Ile Leu Ile Ile
  1185              1190              1195              1200

Asp Glu Ala Thr Ala Asn Val Asp Pro Arg Thr Asp Glu Leu Ile Gln
          1205              1210              1215

Lys Lys Ser Gly Arg Asn Leu Pro Thr Ala Pro Cys
  1220              1225

<210> 538
<211> 1262
<212> PRT
<213> Homo sapiens

<400> 538
Met Tyr Ser Val Leu Pro Glu Asp Arg Ser Gln His Leu Gly Glu Glu

```

5					10					15					
Leu	Gln	Gly	Phe	Trp	Asp	Lys	Glu	Val	Leu	Arg	Ala	Glu	Asn	Asp	Ala
			20					25					30		
Gln	Lys	Pro	Ser	Leu	Thr	Arg	Ala	Ile	Ile	Lys	Cys	Tyr	Trp	Lys	Ser
		35					40					45			
Tyr	Leu	Val	Leu	Gly	Ile	Phe	Thr	Leu	Ile	Glu	Glu	Ser	Ala	Lys	Val
	50					55					60				
Ile	Gln	Pro	Ile	Phe	Leu	Gly	Lys	Ile	Ile	Asn	Tyr	Phe	Glu	Asn	Tyr
65						70					75				80
Asp	Pro	Met	Asp	Ser	Val	Ala	Leu	Asn	Thr	Ala	Tyr	Ala	Tyr	Ala	Thr
				85					90					95	
Val	Leu	Thr	Phe	Cys	Thr	Leu	Ile	Leu	Ala	Ile	Leu	His	His	Leu	Tyr
			100					105					110		
Phe	Tyr	His	Val	Gln	Cys	Ala	Gly	Met	Arg	Leu	Arg	Val	Ala	Met	Cys
		115					120					125			
His	Met	Ile	Tyr	Arg	Lys	Ala	Leu	Arg	Leu	Ser	Asn	Met	Ala	Met	Gly
	130					135					140				
Lys	Thr	Thr	Thr	Gly	Gln	Ile	Val	Asn	Leu	Leu	Ser	Asn	Asp	Val	Asn
145						150					155				160
Lys	Phe	Asp	Gln	Val	Thr	Val	Phe	Leu	His	Phe	Leu	Trp	Ala	Gly	Pro
			165					170						175	
Leu	Gln	Ala	Ile	Ala	Val	Thr	Ala	Leu	Leu	Trp	Met	Glu	Ile	Gly	Ile
			180					185					190		
Ser	Cys	Leu	Ala	Gly	Met	Ala	Val	Leu	Ile	Ile	Leu	Leu	Pro	Leu	Gln
		195					200					205			
Ser	Cys	Phe	Gly	Lys	Leu	Phe	Ser	Ser	Leu	Arg	Ser	Lys	Thr	Ala	Thr
	210					215					220				
Phe	Thr	Asp	Ala	Arg	Ile	Arg	Thr	Met	Asn	Glu	Val	Ile	Thr	Gly	Ile
225						230					235				240
Arg	Ile	Ile	Lys	Met	Tyr	Ala	Trp	Glu	Lys	Ser	Phe	Ser	Asn	Leu	Ile
			245					250						255	
Thr	Asn	Leu	Arg	Lys	Lys	Glu	Ile	Ser	Lys	Ile	Leu	Arg	Ser	Ser	Cys
		260					265					270			
Leu	Arg	Gly	Met	Asn	Leu	Ala	Ser	Phe	Phe	Ser	Ala	Ser	Lys	Ile	Ile

275								280								285
Val	Phe	Val	Thr	Phe	Thr	Thr	Tyr	Val	Leu	Leu	Gly	Ser	Val	Ile	Thr	
290							295				300					
Ala	Ser	Arg	Val	Phe	Val	Ala	Val	Thr	Leu	Tyr	Gly	Ala	Val	Arg	Leu	
305				310						315					320	
Thr	Val	Thr	Leu	Phe	Phe	Pro	Ser	Ala	Ile	Glu	Arg	Val	Ser	Glu	Ala	
				325					330					335		
Ile	Val	Ser	Ile	Arg	Arg	Ile	Gln	Thr	Phe	Leu	Leu	Leu	Asp	Glu	Ile	
			340					345					350			
Ser	Gln	Arg	Asn	Arg	Gln	Leu	Pro	Ser	Asp	Gly	Lys	Lys	Met	Val	His	
		355					360					365				
Val	Gln	Asp	Phe	Thr	Ala	Phe	Trp	Asp	Lys	Ala	Ser	Glu	Thr	Pro	Thr	
	370					375					380					
Leu	Gln	Gly	Leu	Ser	Phe	Thr	Val	Arg	Pro	Gly	Glu	Leu	Leu	Ala	Val	
385					390					395					400	
Val	Gly	Pro	Val	Gly	Ala	Gly	Lys	Ser	Ser	Leu	Leu	Ser	Ala	Val	Leu	
				405					410					415		
Gly	Glu	Leu	Ala	Pro	Ser	His	Gly	Leu	Val	Ser	Val	His	Gly	Arg	Ile	
			420					425					430			
Ala	Tyr	Val	Ser	Gln	Gln	Pro	Trp	Val	Phe	Ser	Gly	Thr	Leu	Arg	Ser	
		435					440					445				
Asn	Ile	Leu	Phe	Gly	Lys	Lys	Tyr	Glu	Lys	Glu	Arg	Tyr	Glu	Lys	Val	
	450					455					460					
Ile	Lys	Ala	Cys	Ala	Leu	Lys	Lys	Asp	Leu	Gln	Leu	Leu	Glu	Asp	Gly	
465					470				475					480		
Asp	Leu	Thr	Val	Ile	Gly	Asp	Arg	Gly	Thr	Thr	Leu	Ser	Gly	Gly	Gln	
				485					490					495		
Lys	Ala	Arg	Val	Asn	Leu	Ala	Arg	Ala	Val	Tyr	Gln	Asp	Ala	Asp	Ile	
			500					505					510			
Tyr	Leu	Leu	Asp	Asp	Pro	Leu	Ser	Ala	Val	Asp	Ala	Glu	Val	Ser	Arg	
	515						520					525				
His	Leu	Phe	Glu	Leu	Cys	Ile	Cys	Gln	Ile	Leu	His	Glu	Lys	Ile	Thr	
	530					535					540					
Ile	Leu	Val	Thr	His	Gln	Leu	Gln	Tyr	Leu	Lys	Ala	Ala	Ser	Gln	Ile	

	820		825		830	
Ile Phe Leu Arg Arg Tyr Phe Leu Glu Thr Ser Arg Asp Val Lys Arg	835		840		845	
Leu Glu Ser Thr Thr Arg Ser Pro Val Phe Ser His Leu Ser Ser Ser	850		855		860	
Leu Gln Gly Leu Trp Thr Ile Arg Ala Tyr Lys Ala Glu Glu Arg Cys	865		870		875	880
Gln Glu Leu Phe Asp Ala His Gln Asp Leu His Ser Glu Ala Trp Phe	885		890		895	
Leu Phe Leu Thr Thr Ser Arg Trp Phe Ala Val Arg Leu Asp Ala Ile	900		905		910	
Cys Ala Met Phe Val Ile Ile Val Ala Phe Gly Ser Leu Ile Leu Ala	915		920		925	
Lys Thr Leu Asp Ala Gly Gln Val Gly Leu Ala Leu Ser Tyr Ala Leu	930		935		940	
Thr Leu Met Gly Met Phe Gln Trp Cys Val Arg Gln Ser Ala Glu Val	945		950		955	960
Glu Asn Met Met Ile Ser Val Glu Arg Val Ile Glu Tyr Thr Asp Leu	965		970		975	
Glu Lys Glu Ala Pro Trp Glu Tyr Gln Lys Arg Pro Pro Pro Ala Trp	980		985		990	
Pro His Glu Gly Val Ile Ile Phe Asp Asn Val Asn Phe Met Tyr Ser	995		1000		1005	
Pro Gly Gly Pro Leu Val Leu Lys His Leu Thr Ala Leu Ile Lys Ser	1010		1015		1020	
Gln Glu Lys Val Gly Ile Val Gly Arg Thr Gly Ala Gly Lys Ser Ser	1025		1030		1035	1040
Leu Ile Ser Ala Leu Phe Arg Leu Ser Glu Pro Glu Gly Lys Ile Trp	1045		1050		1055	
Ile Asp Lys Ile Leu Thr Thr Glu Ile Gly Leu His Asp Leu Arg Lys	1060		1065		1070	
Lys Met Ser Ile Ile Pro Gln Glu Pro Val Leu Phe Thr Gly Thr Met	1075		1080		1085	
Arg Lys Asn Leu Asp Pro Phe Asn Glu His Thr Asp Glu Glu Leu Trp						

1090 1095 1100
 Asn Ala Leu Gln Glu Val Gln Leu Lys Glu Thr Ile Glu Asp Leu Pro
 1105 1110 1115 1120
 Gly Lys Met Asp Thr Glu Leu Ala Glu Ser Gly Ser Asn Phe Ser Val
 1125 1130 1135
 Gly Gln Arg Gln Leu Val Cys Leu Ala Arg Ala Ile Leu Arg Lys Asn
 1140 1145 1150
 Gln Ile Leu Ile Ile Asp Glu Ala Thr Ala Asn Val Asp Pro Arg Thr
 1155 1160 1165
 Asp Glu Leu Ile Gln Lys Lys Ile Arg Glu Lys Phe Ala His Cys Thr
 1170 1175 1180
 Val Leu Thr Ile Ala His Arg Leu Asn Thr Ile Ile Asp Ser Asp Lys
 1185 1190 1195 1200
 Ile Met Val Leu Asp Ser Gly Arg Leu Lys Glu Tyr Asp Glu Pro Tyr
 1205 1210 1215
 Val Leu Leu Gln Asn Lys Glu Ser Leu Phe Tyr Lys Met Val Gln Gln
 1220 1225 1230
 Leu Gly Lys Ala Glu Ala Ala Ala Leu Thr Glu Thr Ala Lys Gln Arg
 1235 1240 1245
 Trp Gly Phe Thr Met Leu Ala Arg Leu Val Ser Asn Ser
 1250 1255 1260
 <210> 539
 <211> 10
 <212> PRT
 <213> Artificial Sequence

 <220>
 <223> Made in a lab

 <400> 539
 Cys Leu Ser His Ser Val Ala Val Val Thr
 1 5 10

 <210> 540
 <211> 9
 <212> PRT
 <213> Artificial Sequence

 <220>
 <223> Made in a lab

002290-0343360

<400> 540

Ala Val Val Thr Ala Ser Ala Ala Leu

1 5

<210> 541

<211> 14

<212> PRT

<213> Homo sapiens

<400> 541

Leu Ala Gly Leu Leu Cys Pro Asp Pro Arg Pro Leu Glu Leu

5 10

<210> 542

<211> 15

<212> PRT

<213> Homo sapiens

<400> 542

Thr Gln Val Val Phe Asp Lys Ser Asp Leu Ala Lys Tyr Ser Ala

5 10 15

<210> 543

<211> 12

<212> PRT

<213> Homo sapiens

<400> 543

Phe Met Gly Ser Ile Val Gln Leu Ser Gln Ser Val

5 10

<210> 544

<211> 18

<212> PRT

<213> Homo sapiens

<400> 544

Thr Tyr Val Pro Pro Leu Leu Leu Glu Val Gly Val Glu Glu Lys Phe

5 10 15

Met Thr

<210> 545

<211> 18

<212> PRT

Ser Val

Thr Glu Ala Arg Arg His Tyr Asp Glu Gly Val Arg Met
20 25

Ser Ala Pro Ser Leu Ser Pro His Cys Cys Pro Cys Arg Ala Arg Leu
20 25 30

Ala Phe Arg Asn Leu Gly Ala Leu Leu Pro Arg Leu His Gln Leu Cys
35 40 45

Cys Arg Met Pro Arg Thr Leu Arg Arg Leu
50 55

Ile Asp Trp Asp Thr Ser Ala Leu Ala Pro Tyr Leu Gly Thr Gln Glu
5 10 15

```
<400> 552
agcatatgta acatgacctg tgcttcagtg ttcttttgtg atcaaaaatt ccttactttt 60
agttttttat ctatggtaga accaccaga gcaggggtcc tcaactccca ggccacagac 120
tcataccagt ccacggacta ttatgaacca caccacacag gaggaggtga gcactaggca 180
agccaaggaa gcttcacctg tacttacagc cacacgccat ggctcatatt acagcctgaa 240
ctctgcctcc actcagatca gtgataacat tagaaaactca ttggagcacg aacctgtttg 300
tgaactgcct atccgaagga tctaggttgt gtgcttcgta tgagaatcta atgccagatg 360
```

```

atctatcatt gtctcacttt gccccagat aagaccatct agttgcagaa aaataagctc 420
agagcttcca ctgattctac attatggata tgtgccgcgc aagcaagcac aaagccctac 480
ttttacacat gcctagtgat gcttcatgga caaggcttgg ctctgttgag tccaactaac 540
ctacctgaga ttctgagatt tctcttcaat ggcttctgtg gagctagagt ttgaaaatat 600
cttaaaatct tgagctagag atggaagtag cttggacgat tttcattatc atgtaaatcg 660
ggtcactcaa ggggcccaacc acagctggga gccactgctc aggggaaggt tcatatggga 720
ctttctactg cccaaggttc tatacaggat ataaagggtgc ctcacagtat agatctggta 780
gcaaagaaga agaaacaaac actgatctct ttctgccacc cctctgacct tttggaactc 840
ctctgacctt ttagaacaag cctacctaat atctgctaga gaaaagacca acaacggcct 900
caaaggatct cttaccatga aggtctcagc taattcttgg ctaagatgtg gggtccacat 960
taggttctga atatgggggg aagggtcaat ttgctcattt tgtgtgtgga taaagtcagg 1020
atgccagagg gccagagcag ggggctgctg ctttggaac aatggctgag catataacca 1080
taggtatggg aacaaaaaac atcaaagtca ctgtatcaat tgccatgaag actcgaggga 1140
cctgaatcta ccgattcatc ttaaggcagc aggaccagtt tgagtggcaa caatgcagca 1200
gcagaatcaa tggaaacaac agaagtattg caatgtcctt ttttttctcc tccttctgac 1260
ttgataaaag ggaccgtctt ccttggattt agtgaacccc tttggttcct gaaaaattca 1320
aggagtatct aggacatagt cccagaaga cagtacaaga ctttctgata aactggacat 1380
ttcaagcccc aaataactaa tcagaaaaat caaagatgtg atactatttt ttatcccatg 1440
cataggtgct acacttggat caaatgaaca atgttgggat ctytatggat aaaggtctta 1500
aaagtcttga gataaagaat cctgcacca ctggtacttc taacttgtct tgttttttgt 1560
ctgwtttctg gctgatgcag gggactaact cactgccacg cgaaaactac ctgaactgaa 1620
ctatgacatc tcacctgata tgtaagatgt aactgttata attattttta acctcaattt 1680
agcattaact agccttttaa tgtaaacact tacacattat gaygactaga aacagcatc 1740
tctctggccg tctgtccaga tagatcttga gaagatacat caatgttttg ctcaagtaga 1800
aggctgacta tacttgcgga tccacaacat acagcaagta tgagagcagt tctaaaatga 1860
cagagatagg aacagtaata aagttattkt aaaagcta attgatatact ttaccaattt 1920
aacatcttgc ctgtccgtgc agaatcaaac atttacatgc actaaaagac ataagcatct 1980
tcagtgtctc agtgttcatc tttgtaaaat accaccaagg ttaaaaggaa gggacaaaaa 2040
aaaaaaaccc tcttatctca gtggggtatt gcatagcaga agctactaat ttgaagtcct 2100
ttgatggaca agaaacaata ttaggggccac ttatctgaaa tgaacaaaga tttagttaa 2160
gatttcatca cagcttccct agactgatat gctgtaatag aaaatcagct agggggtaaa 2220
ataaataaga gctctctgca tgctgaaagc aagtaagatt aataataatg gtaagaatag 2280
tagtcacagg agtttcagtt aatgatgcca ataagcatgt gctaggcact gaattaaatg 2340
ccacatatat ctttcttatg cgcagcaaac tttgaaggat atattctcct acttttcata 2400
tatgacaaca tatttggtgg taaataacgt tccaagggtc acacacctag caagtaagaa 2460
agttaggaat taaacccagt attgtgtgaa tctaaagcct aacttttttc tctttatcac 2520
ccacctacgg cttgtcttca ttaaaggaaa agtgtatcca cttaaaaaaa aaaaaaa 2577

```

<210> 553

<211> 58

<212> PRT

<213> Homo sapiens

<400> 553

```

Ser Ile Cys Asn Met Thr Cys Ala Ser Val Phe Phe Cys Asp Gln Lys
          5                      10                      15

```

```

Phe Leu Thr Phe Ser Phe Leu Ser Met Val Glu Pro Pro Arg Ala Gly
          20                      25                      30

```

```

Val Leu Asn Ser Gln Ala Thr Asp Ser Tyr Gln Ser Thr Asp Tyr Tyr

```

45

<213> Homo sapiens

<400> 556

Asn His Pro Glu Gln Gly Ser Ser Thr Pro Arg Pro Gln Thr His Thr
 5 10 15

Ser Pro Arg Thr Ile Met Asn His Thr Thr Gln Glu Glu Val Ser Thr
 20 25 30

Arg Gln Ala Lys Glu Ala Ser Pro Val Leu Thr Ala Thr Arg His Gly
 35 40 45

Ser Tyr Tyr Ser Leu Asn Ser Ala Ser Thr Gln Ile Ser Asp Asn Ile
 50 55 60

Arg Asn Ser Leu Glu His Glu Pro Cys Cys Glu Leu Pro Ile Arg Arg
 65 70 75 80

Ile

<210> 557

<211> 54

<212> PRT

<213> Homo sapiens

<400> 557

Ser Leu Ser Ala Thr Pro Leu Thr Leu Trp Asn Ser Ser Asp Pro Leu
 5 10 15

Glu Gln Ala Tyr Leu Ile Ser Ala Arg Glu Lys Thr Asn Asn Gly Leu
 20 25 30

Lys Gly Ser Leu Thr Met Lys Val Ser Ala Asn Ser Trp Leu Arg Cys
 35 40 45

Gly Phe His Ile Arg Phe
 50

<210> 558

<211> 77

<212> PRT

<213> Homo sapiens

<220>

<221> VARIANT

<222> (1)...(77)

<223> Xaa = Any amino acid

<400> 558

55

<213> Homo sapiens

<223> Xaa = Any amino acid

<213> Homo sapiens

<223> Xaa = Any amino acid

Leu Ser Ser Gly Asp Tyr Val Leu Asp Thr Pro
50 55

<210> 563
 <211> 79
 <212> PRT
 <213> Homo sapiens

<400> 563
 Cys Phe Leu Phe Pro Tyr Leu Trp Leu Tyr Ala Gln Pro Leu Phe Pro
 5 10 15
 Lys Gln Gln Pro Pro Ala Leu Ala Pro Gly His Pro Asp Phe Ile His
 20 25 30
 Thr Gln Asn Glu Gln Ile Asp Pro Ser Pro His Ile Gln Asn Leu Met
 35 40 45
 Trp Asn Pro His Leu Ser Gln Glu Leu Ala Glu Thr Phe Met Val Arg
 50 55 60
 Asp Pro Leu Arg Pro Leu Leu Val Phe Ser Leu Ala Asp Ile Arg
 65 70 75

<210> 564
 <211> 64
 <212> PRT
 <213> Homo sapiens

<400> 564
 Ala Cys Ser Lys Gly Ser Glu Glu Phe Gln Arg Val Arg Gly Val Ala
 5 10 15
 Glu Arg Asp Gln Cys Leu Phe Leu Leu Leu Cys Tyr Gln Ile Tyr Thr
 20 25 30
 Val Arg His Leu Tyr Ile Leu Tyr Arg Thr Leu Gly Ser Arg Lys Ser
 35 40 45
 His Met Asn Leu Pro Leu Ser Ser Gly Ser Gln Leu Trp Leu Ala Pro
 50 55 60

<210> 565
 <211> 57
 <212> PRT
 <213> Homo sapiens

<220>
 <221> VARIANT
 <222> (1)...(57)
 <223> Xaa = Any amino acid

<400> 565

Leu Tyr Tyr Cys Ser Tyr Leu Cys His Phe Arg Thr Ala Leu Ile Leu
 5 10 15

Ala Val Cys Cys Gly Ser Ala Ser Ile Val Ser Leu Leu Leu Glu Gln
 20 25 30

Asn Ile Asp Val Ser Ser Gln Asp Leu Ser Gly Gln Thr Ala Arg Glu
 35 40 45

Tyr Ala Val Ser Ser Xaa His Asn Val
 50 55

<210> 566

<211> 55

<212> PRT

<213> Homo sapiens

<400> 566

Ile Leu Leu Glu Phe Phe Arg Asn Gln Arg Gly Ser Leu Asn Pro Arg
 5 10 15

Lys Thr Val Pro Phe Ile Lys Ser Glu Gly Gly Glu Lys Lys Gly His
 20 25 30

Cys Asn His Ser Val Val Ser Ile Asp Ser Ala Ala Ala Leu Leu Pro
 35 40 45

Leu Lys Leu Val Leu Leu Pro
 50 55

<210> 567

<211> 51

<212> PRT

<213> Homo sapiens

<400> 567

Tyr Ser Asp Phe Asp Val Phe Cys Ser His Thr Tyr Gly Tyr Met Leu
 5 10 15

Ser His Cys Ser Gln Ser Ser Ser Pro Leu Leu Trp Pro Leu Gly Ile
 20 25 30

Leu Thr Leu Ser Thr His Lys Met Ser Lys Leu Thr Leu Pro Pro Ile
 35 40 45

Phe Arg Thr
 50

<213> Homo sapiens

Phe Leu Leu Gln His Ile Ser Leu Gly Lys Leu
65 70 75

<213> Homo sapiens

gcatccagag	tgttgactg	gttacaggct	atgaacctac	actgatgcgg	caccaccacc	60
cagagtccac	rggttatgtt	ggttcacatt	tactcttgct	gtggtatggt	ctatagggtt	120
ggacagatgt	ccgataatcc	tttttacatt	ttggcactct	tgggtagctc	gtcttgtagg	180
aatggacttg	cttcaaagtg	gaggcaggca	gatacttcag	acgggtatat	ggagccctgt	240
tttcagttgc	ttttctaatt	ctctcttata	gtttacctca	aaatcttctc	gagggtctcg	300
ttctttttaa	aatctttgtc	tactttgcag	catcactctg	acactcccat	tgattctctca	360
gcacctactg	actacacggt	taggagtgc	agggtagaat	tcatgtttta	ttcatctttg	420
ggtctgtagc	accagcaaa	gtgctcagta	aatgcgcagt	aattgatttg	acctctgaac	480
aaatacacac	tgtactaaga	atctacacac	cgaagacaa	aaacaagaca	aatttgagtg	540
ctacaggtgt	cacgcttggc	atcacacatg	tgcctgtgta	ttctctaggt	tggttaccag	600
gagctctgcc	actgcatgtc	cactagtgc	gggttcgctc	caccacccca	gctgggtagc	660
cgtgctctc	acataagggg	tccaattaaa	attgccagga	ataaattccc	ccggactttg	720
acttctcaag	agctaagaag	gtttgctgag	tattctggca	tgatgtttgg	tgatcaaaca	780
actgtgggcc	aaaaatgatg	agtatttccc	cctcttgctg	aagatgtgct	ccatacaata	840
gtccatcaca	ttcatcattc	atcagtctgg	aagtgtgcag	aacaacatgt	aatagataat	900
atgattggct	gcacacttcc	agactgatga	atgatgaatg	tgatggacta	ttgtatggag	960
cacatottca	gcaagagggg	gaaatactca	tcatttttata	tattacatgt	tgttctgggt	1020
tttttttttt	tccaatgtcc	agcctaaact	ataaagtact	ttgagaacgc	acagtgaacc	1080
ataagottgc	caataaagag	tcctctgtgg	tatggaaactg	gcttattttca	tacacaatct	1140
gcaaacaatg	agggcactat	tggaaacata	ctgtgctgca	cagagcattt	acaccgctta	1200
tctttaatct	tcccagcaa	tccttgcttt	gtgcgcattt	atgatccttg	ctctcagaag	1260

tccacatact tttccccaac cgtaacaaat tattttaactc atctaagtga tgtatgtccg 1320
 cgcagtctga aaacagtaat tgtccttggg aagaagtgaag ttttaagagag ctctagggca 1380
 ctcatcaciaa ctccagccct gccctccatg tggtagcagc tctttggact ggggctaagt 1440
 gcttattctt gtgcttcatt cctggtaagc tcaatttctt taccttagga taactttgct 1500
 ggaaaagggc tcagattcag ccgaccattg tggcctctgt ggctgtcaca gcttgtccct 1560
 gacatgctat gatgttgggt ccccttctca tcccttggg atttcttctg ctggcccaca 1620
 gccagaacaa ctaggccttt tactccacca tcccttgggt tcttttgggt tcgttggtaa 1680
 aaatcaatcc ttctaccatc catgcatagc aatttctaaa aactgaattt caagagcagt 1740
 atctgaagaa acaaacatga tttggctcct ttagtaaaca gaataaattt taataaatca 1800
 actttgaaat agttgtaaga gttaagaaaa agcacaaaac tgagatcatc agagcagctt 1860
 ggctcaaag gacaggcagc aggattctac aggggttgag ccttcttaag tgaagctgtt 1920
 tcctgcaggc tcctgtctcc aagctcctag ctaacagccc cttctcccac gattggcaac 1980
 aaagagcaaa aataactttg tacttgatgc tgagtcagtg taaaaagcca taaaaaattc 2040
 cctctaaatg tcaaaatggt tgctcctttt gaggtctctc tctctctact ggggtctggat 2100
 aaattagcac tgggcttata ttgagtcaca gatctggggt ctgccacaga gagcttctct 2160
 ctagtgtgtg atgctttttc tccaaactat tgatacaaaa tgcactggaa tagaaatcaa 2220
 cagaaactgg tcaaaaggtg ggcatacaca ttctcatgta gatgtaaagc tgtgcttaga 2280
 attcctttgt ggagctctgg ttggtcttgg ttttcttgggt gtttgattca tttttttacg 2340
 taaattacaa aaaccctcca catttcttca tggattgtat tagtccatgt tctccagaga 2400
 agcagaacga gttggatgta tgttttggaa gagattatga ggaaccgggt catgtgatga 2460
 aggaggttga gaggtcctgt gctctgccat ctgcaagctg aagacctgga aagctgaggg 2520
 tgtggctcca gtctgagctc gaaggcccaa gaaccagggg aaccaacggg ttagattcca 2580
 ggttgaaggc aggagaagat ggatgtccca gctcagcagg caggcaggaa gcaaatgggg 2640
 taaattcctc cttcctccac cttttgttcc attcaggcct tcaacagatt ggatgagcgc 2700
 cccccaccc ccacactagg gagggccatc tgctttactg agtcggctga gtcaagtgcc 2760
 agcctcatcc caaaacactc tccagacaca cgcagaaatg tttcatctgg gcacctgtg 2820
 gccagtcagc ctgacacaca gaactaacca tgacatggat tcttcttaaa gcagtgatag 2880
 gagcgaacag aaacattttc ataattttca attattttta atgaaaacta tatctgatgg 2940
 aattgtttaa acctagtctg gccacacatt atttcttggg accgcccctc cttcaatccc 3000
 ttggacactg atgactttat gccagatta cactggaggc ctgtgctgat tttctaacac 3060
 atacctgcaa ctgagctggc aaaaagaaaa ctaggcaagt atgacagata catgatgcac 3120
 aggctaagtg caaaggaaaag aaaaacacca actgcaggga tgagggactc accccttag 3180
 aagtttctac ttgagcagct agaagactac aatgccactc atcaaacag tgactcaggg 3240
 ggagtatttg ggataaagga ggaatctgat gttggagggt aaatttgaag tgtctttaag 3300
 acctacaggt aacgagacag ctggacaaac acatggaact caggacaaag gctctaagga 3360
 cagcacagca gctgacatcc tgtgtgacag ccttgaaagc agcaggcccg ccgctcacat 3420
 tttggaaggg aaaatgggtg caatgttgtc tgccactttg gggccttctt gggtcacatg 3480
 cattttacat ttatgcagtt gatatattta tgtttcctgg gtcttttata cattagacac 3540
 catgattctc aatcctttgt tattttgtat taaaaaagc tgaattatta tttcaaatat 3600
 gggcaaatta gagccttcca tattgccaag gtgtatcaac cacactgata ycaygatctc 3660
 tcttttgaat tagttttcca gtccacacct accatttatt tcatgattgg tttcagactt 3720
 gttcctcctg gaaacactcc ctaacaagca cccttgccagg aatgaagaca caccacacac 3780
 atctacccca ttactgcatg tactcaagag tcagctttta tatgatctct cccaagtgtc 3840
 cctataatgg ggatctttca ctccacctaa agtgaggaca aaataactga aagcatgagc 3900
 ccagtgcctg taggtgtgca attaacctca gaccaaggaa gtgccgaacg catctggctt 3960
 ttagcaaggc acctgacaaa gtoccttcagg atgtttttgt acatgagcta gagaaatgta 4020
 cctggagaac agcttctact gccagatgat ctactcaaa agatgcagat taagcaaaat 4080
 atcaacccaa aggggtgtcc ctgatggccc accagcccct gtgcctgggt cgtttcctat 4140
 gtttcttaga tttggtttca gaactgtctc tctgcagac actccctaac cagcatcctt 4200
 gcagaaaact ggtgaactag aaaaggcctg tgtgggtcac gtggccaccc aacaccacag 4260
 cagtgtctaa ggtatgcgtg ggagcctgca cagcaggagc ggggtcttct ggagacccgc 4320

```

atgagatgca aagggcagtg gacaaggagc caagggaggt ggctctagtc acgctggtat 4380
gggtgccagct tgaggatgct gggcaagtcc cgagccgtct gccttcctag taccacagtt 4440
accactgtct gttacctcgc gagttcaagt gcttcacgtg agacagctac gagacaggcc 4500
cctggaaact ggaaaatgcy aagtaaagt catgcacaat tgttggtcac attttatctc 4560
aatcactttt accaaatcag gctaaaccct gggattcat aacgtcttg gctgtacaaa 4620
ttgttccttg aaatgactca gagacatttt ctgaattggc ttccatcagc caagcatttc 4680
ttcagaactg gaaaaatgct ttaaatttgg ctttgtcatg attattaaaa cactctgtac 4740
atTTTTtatt attgaaatta acacattgcc tactttttaa aaattggaaa aagaaaaaaa 4800
aaaaaaaaa 4809

```

<210> 570

<211> 951

<212> DNA

<213> Homo sapiens

<400> 570

```

aaaattgaat attgagatac cattcttttag tgttaccttt tttaccacaca tgtgtttctg 60
aaaatattgg aattttattc atcttaaaaa ttggaccogg ccttattttac catctttaat 120
ccatttttagt actatgggtg agtacatgga attgaagtct ggcttaaate ttcagaaagt 180
tatatatcta ttttatttta tttttttgag acagagtctc gctgtgtcac ccaggctgga 240
gtgcggtgcc acaatcttgg ctactgcaa cctctgagtc ccaggttcaa gcgatactca 300
tgccctcgcc tctlgagtag ctgggactac aggcgtgcac caccacatct ggctaattct 360
tttttgtatt tttagtagag acgggggttc actgtggtct ccctctcccg acctcgtgat 420
ccgcctgcct cccaaagtgc tgggattaca ggcattgagc accgcacaca gctgggactg 480
ggtaatttat aaagaaaaga ggtttaatga ctacagttc cgcatggctg gagaggcctc 540
aggaaactca caatcatggt ggaaggcgaa ggggaagcaa ggcacgtctt acatggtggc 600
aggagagaac gagtgagggg ggagactgcc acaaactttt tttttttgag acaagagtct 660
ggccctgttg cccaggctgg agtgcagtgg catgatctca gctcactgca acctctgcct 720
cacaggttca agcaattctc atgcctcagc ctccgcgata gctgggacca caggatgca 780
ccaccacacc tagctaattt ttgtagtttt agtagagatg gggctctcact atgttgctca 840
ggctggtcta aaactcctgg gctccagcaa tcgcctgcc ttggcctccc aaagtgctgg 900
ggttacaggc ataagccacc acatccagcc tgccacatac ttttaaacta t 951

```

<210> 571

<211> 819

<212> DNA

<213> Homo sapiens

<400> 571

```

cagcttaaaa atggtttctt gaaatcagtg attagcattc actcaccagt acccctacta 60
aggggtaggc actggtttgt actcctggga atacaggagt acaccagaat ttatttctgc 120
ttattgcttt tgttgcaaat gccgtggctt catctgagga attctagaat tcagagggtg 180
tagccctcca ctctgctgtc ttgctatctg ctctcattgc atccgtttta cctgcattct 240
gaaagatggt tctcaggttt ttcttgaag attttcttct tttctgattc tgacaatggt 300
ttaaatcatt gtactgtggg tatcatttct ctgcatttat tttaccatc ttcttttgta 360
acttgtccta ttgtctttta atttctgct gttctttatg gctttcaact tcataaataa 420
catgttttct caaatctctt tgtgaattcc agagagggcc aggcacggtg gctcacatct 480
gtaatcccag cactttgggg aggctgagac ggggtggatca cttgaggtca ggagtttgag 540
accagcctgg ccaacatggt gaaatcccg ttcactaaaa atacaaaaat taccaggca 600
tggtggcggg cgcctgtaat cccaggctct cgggaggctg agggaggaga atcgcttgaa 660
cctgggaggg tgaggggagga gaatcgcttg aaccggggag gcagagggtg cagtgaaccg 720

```

```
<210> 574
<211> 63
<212> PRT
```

<213> Homo sapiens

<400> 574

```
Met Thr His Ser Ser Ala Trp Leu Glu Arg Pro Gln Glu Thr Tyr Asn
              5              10              15

His Gly Gly Arg Arg Arg Gly Ser Lys Ala Arg Leu Thr Trp Trp Gln
              20              25              30

Glu Arg Thr Ser Glu Gly Gly Asp Cys His Lys Leu Phe Phe Phe Glu
              35              40              45

Thr Arg Val Trp Pro Cys Cys Pro Gly Trp Ser Ala Val Ala
              50              55              60
```

<210> 575

<211> 77

<212> PRT

<213> Homo sapiens

<400> 575

```
Met Val Lys Ser Arg Phe Thr Lys Asn Thr Lys Ile Thr Gln Ala Trp
              5              10              15

Trp Arg Ala Pro Val Ile Pro Gly Thr Arg Glu Ala Glu Gly Gly Glu
              20              25              30

Ser Leu Glu Pro Gly Arg Leu Arg Glu Glu Asn Arg Leu Asn Pro Gly
              35              40              45

Gly Arg Gly Cys Ser Glu Pro Arg Ser Cys Cys Cys Thr Pro Ala Trp
              50              55              60

Ser Thr Glu Gln Asp Ser Ala Ser Lys Thr Asn Lys
              65              70              75
```

<210> 576

<211> 69

<212> PRT

<213> Homo sapiens

<220>

<221> unsure

<222> (42)

<223> Xaa = Any Amino Acid

<400> 576

```
Met Leu Gly Lys Ser Arg Ala Val Cys Leu Pro Ser Thr Thr Val Thr
              5              10              15
```

Gln Pro His
50

<210> 579

<211> 57

<212> PRT

<213> Homo sapiens

<400> 579

Met His Phe Thr Phe Met Gln Leu Ile Tyr Leu Cys Phe Leu Gly Leu
 5 10 15

Leu Tyr Ile Arg His His Asp Ser Gln Ser Phe Val Ile Leu Tyr Tyr
 20 25 30

Lys Lys Leu Asn Tyr Tyr Phe Lys Tyr Gly Gln Ile Arg Ala Phe His
 35 40 45

Ile Ala Lys Val Tyr Gln Pro His
 50 55

<210> 580

<211> 68

<212> PRT

<213> Homo sapiens

<400> 580

Met Glu Leu Arg Thr Lys Ala Leu Arg Thr Ala Gln Gln Leu Thr Ser
 5 10 15

Cys Val Thr Ala Leu Lys Ala Ala Gly Pro Pro Leu Thr Phe Trp Lys
 20 25 30

Gly Lys Trp Val Gln Cys Cys Leu Pro Leu Trp Gly Leu Leu Gly Ser
 35 40 45

His Ala Phe Tyr Ile Tyr Ala Val Asp Ile Phe Met Phe Pro Gly Ser
 50 55 60

Phe Ile His
 65

<210> 581

<211> 78

<212> PRT

<213> Homo sapiens

<400> 581

Met Leu Glu Val Lys Phe Glu Val Ser Leu Arg Pro Thr Gly Asn Glu
 5 10 15

Met Met Phe Gly Asp Gln Thr Thr Ala Gly Gln Lys
50 55 60

<211> 77

<213> Homo sapiens

Met Cys Leu Cys Ile Pro Leu Gly Gly Tyr Gln Glu Leu Cys His Cys
5 10 15

Met Ser Thr Ser Asp Gly Phe Ala Pro Pro Pro Gln Leu Gly Ser Arg
20 25 30

Cys Ser His Ile Arg Gly Pro Ile Lys Ile Ala Arg Asn Lys Phe Pro
35 40 45

Arg Thr Leu Thr Ser Gln Glu Leu Arg Arg Phe Ala Glu Tyr Ser Gly
50 55 60

Met Met Phe Gly Asp Gln Thr Thr Ala Gly Gln Lys
65 70 75

<211> 51

<213> Homo sapiens

Met Val Tyr Arg Phe Gly Gln Met Ser Asp Asn Pro Phe Tyr Ile Leu
5 10 15

Ala Ser Leu Gly Ser Ser Ser Cys Arg Asn Gly Leu Ala Ser Lys Trp
20 25 30

Arg Gln Ala Asp Pro Ser Asp Gly Tyr Met Glu Pro Cys Phe Gln Leu
35 40 45

Leu Phe
50

<211> 61

<213> Homo sapiens

Met Leu Val His Ile Tyr Ser Cys Cys Gly Met Val Tyr Arg Phe Gly
5 10 15

Gln Met Ser Asp Asn Pro Phe Tyr Ile Leu Ala Ser Leu Gly Ser Ser

```
<210> 588  
<211> 81  
<212> PRT  
<213> Homo sapiens  
  
<400> 588  
Met Pro Gln Lys Gln Gln Asn Ser Gln Thr Glu Ala Lys Tyr Arg Ala  
                    5                      10                     15  
Leu Gln Phe Arg Gln Tyr Asn Lys Ser Val His Glu Val Asn Leu Lys  
                20                   25                  30
```

Gly Ala Cys Phe Thr Val Ala Gly Leu Pro Arg Ala Trp Thr Thr Gln
 35 40 45

Tyr Ser Ile Ile Asp Lys Arg Ile Arg Gln Glu Ile Tyr Thr Cys Cys
 50 55 60

Leu Ala Phe Val Val Ile Tyr Thr Asn Glu Asn Met Tyr Tyr Ser Tyr
 65 70 75 80

Ile

<210> 589

<211> 157

<212> PRT

<213> Homo sapiens

<400> 589

Met Thr Met Cys Leu Cys Val Ala Pro Met Gly Arg Ala Thr Arg Met
 5 10 15

Ser Val Thr Cys Asp Arg Leu His Ala Asn Ser Arg Val Arg Tyr Leu
 20 25 30

Trp Cys Gln Lys Asp His Val Pro Gln Met Gln Asp Gln Asp Leu Glu
 35 40 45

Met Glu Ser Met Lys Ala Leu Glu Lys Leu Val Lys Arg Arg His Pro
 50 55 60

Pro Val Ile Phe Ala Ser Leu Val Gln Asn Val Thr Lys Met Pro Arg
 65 70 75 80

Met Ser Gly Val Cys Val Ile Leu Thr Val Leu Lys Pro Thr Ser Ile
 85 90 95

Pro Ser Ala Leu Leu Met Gly Asn Leu Met Ile Met His Ala Lys Ser
 100 105 110

Lys Lys His Arg Val Arg Asn Arg Arg Lys Leu Lys Ser Cys Leu Trp
 115 120 125

Val Asp Val Lys Ile Thr Gln Leu Gln Leu Leu Ser Leu Lys Met Gly
 130 135 140

Ile Met Gln Glu Gln Ile Met Gln Arg Met Leu Thr Asn
 145 150 155

<210> 590

<211> 347

<212> PRT

<213> Homo sapiens

<400> 590

Met Leu Leu Ile Val Ala Arg Pro Val Lys Leu Ala Ala Phe Pro Thr
 5 10 15

Ser Leu Ser Asp Cys Gln Thr Pro Thr Gly Trp Asn Cys Ser Gly Tyr
 20 25 30

Asp Asp Arg Glu Asn Asp Leu Phe Leu Cys Asp Thr Asn Thr Cys Lys
 35 40 45

Phe Asp Gly Glu Cys Leu Arg Ile Gly Asp Thr Val Thr Cys Val Cys
 50 55 60

Gln Phe Lys Cys Asn Asn Asp Tyr Val Pro Val Cys Gly Ser Asn Gly
 65 70 75 80

Glu Ser Tyr Gln Asn Glu Cys Tyr Leu Arg Gln Ala Ala Cys Lys Gln
 85 90 95

Gln Ser Glu Ile Leu Val Val Ser Glu Gly Ser Cys Ala Thr Asp Ala
 100 105 110

Gly Ser Gly Ser Gly Asp Gly Val His Glu Gly Ser Gly Glu Thr Ser
 115 120 125

Gln Lys Glu Thr Ser Thr Cys Asp Ile Cys Gln Phe Gly Ala Glu Cys
 130 135 140

Asp Glu Asp Ala Glu Asp Val Trp Cys Val Cys Asn Ile Asp Cys Ser
 145 150 155 160

Gln Thr Asn Phe Asn Pro Leu Cys Ala Ser Asp Gly Lys Ser Tyr Asp
 165 170 175

Asn Ala Cys Gln Ile Lys Glu Ala Ser Cys Gln Lys Gln Glu Lys Ile
 180 185 190

Glu Val Met Ser Leu Gly Arg Cys Gln Asp Asn Thr Thr Thr Thr Thr
 195 200 205

Lys Ser Glu Asp Gly His Tyr Ala Arg Thr Asp Tyr Ala Glu Asn Ala
 210 215 220

Asn Lys Leu Glu Glu Ser Ala Arg Glu His His Ile Pro Cys Pro Glu
 225 230 235 240

His Tyr Asn Gly Phe Cys Met His Gly Lys Cys Glu His Ser Ile Asn
 245 250 255

Met Gln Glu Pro Ser Cys Arg Cys Asp Ala Gly Tyr Thr Gly Gln His
 260 265 270

Cys Glu Lys Lys Asp Tyr Ser Val Leu Tyr Val Val Pro Gly Pro Val
 275 280 285

Arg Phe Gln Tyr Val Leu Ile Ala Ala Val Ile Gly Thr Ile Gln Ile
 290 295 300

Ala Val Ile Cys Val Val Val Leu Cys Ile Thr Arg Lys Cys Pro Arg
 305 310 315 320

Ser Asn Arg Ile His Arg Gln Lys Gln Asn Thr Gly His Tyr Ser Ser
 325 330 335

Asp Asn Thr Thr Arg Ala Ser Thr Arg Leu Ile
 340 345

<210> 591

<211> 565

<212> DNA

<213> Homo sapien

<400> 591

actaaagcaa	atgaacaagc	tgacttgcta	gtatcatctg	cattcattga	agcacaagaa	60
cttcatgcct	tgactcatgt	aaatgcaata	ggattaaaaa	ataaatttga	tatcacatgg	120
aaacagacaa	aaaatattgt	acaacattgc	accagtgctc	agattctaca	cctggccact	180
caggaagcaa	gagttaatcc	cagaggtcta	tgtcctaatt	tgttatggca	aatggatgct	240
atgcacgtac	cttcatttgg	aaaattgtca	tttgtccatg	tgacagttga	tacttattca	300
catttcatat	gggcaacctg	ccagacagga	gaaagtactt	cccatgttaa	aagacattta	360
ttatcttggt	ttcctgtcat	gggagttcca	gaaaaagtta	aaacagacaa	tgggccaggt	420
tactgtagta	aagcatttca	aaaattctta	aatcagtgga	aaattacaca	tacaatagga	480
attctctata	attcccaagg	acaggccata	attgaaggaa	ctaatagaac	actcaaagct	540
caattgggta	aacaaaaaaa	aaaaa				565

<210> 592

<211> 188

<212> PRT

<213> Homo sapien

<400> 592

Thr	Lys	Ala	Asn	Glu	Gln	Ala	Asp	Leu	Leu	Val	Ser	Ser	Ala	Phe	Ile
1			5					10					15		
Glu	Ala	Gln	Glu	Leu	His	Ala	Leu	Thr	His	Val	Asn	Ala	Ile	Gly	Leu
			20				25				30				
Lys	Asn	Lys	Phe	Asp	Ile	Thr	Trp	Lys	Gln	Thr	Lys	Asn	Ile	Val	Gln

```

      35              40              45
His Cys Thr Gln Cys Gln Ile Leu His Leu Ala Thr Gln Glu Ala Arg
      50              55              60
Val Asn Pro Arg Gly Leu Cys Pro Asn Val Leu Trp Gln Met Asp Val
      65              70              75              80
Met His Val Pro Ser Phe Gly Lys Leu Ser Phe Val His Val Thr Val
      85              90              95
Asp Thr Tyr Ser His Phe Ile Trp Ala Thr Cys Gln Thr Gly Glu Ser
      100             105             110
Thr Ser His Val Lys Arg His Leu Leu Ser Cys Phe Pro Val Met Gly
      115             120             125
Val Pro Glu Lys Val Lys Thr Asp Asn Gly Pro Gly Tyr Cys Ser Lys
      130             135             140
Ala Phe Gln Lys Phe Leu Asn Gln Trp Lys Ile Thr His Thr Ile Gly
      145             150             155             160
Ile Leu Tyr Asn Ser Gln Gly Gln Ala Ile Ile Glu Gly Thr Asn Arg
      165             170             175
Thr Leu Lys Ala Gln Leu Val Lys Gln Lys Lys Lys
      180             185

```

<210> 593

<211> 271

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(271)

<223> n = A,T,C or G

<400> 593

```

actttatgtt cnagtgcana aancncctg gattgccacc ntactctcag ggctgtgant      60
tgtgcnccca nagcaacctg ggcacgcggg gacagggggg ccnacaattg agggagcgg      120
gtccctagct ggggtctata catgncnggg naagggcngc tgagtnccat nagcaaagga      180
nctagnatnt gcgggggtgc ggccctgggccc tacccttttna agcatccntn gatccactcc      240
angaanccng gggtagncag gtttnccaac a                                     271

```

<210> 594

<211> 376

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(376)

<223> n = A,T,C or G

<400> 594

```

cctttggggg nggggggaac ctttaccatt gtncctcttt atttcatttg gttnggggttc      60
gcgcctcnnn gggccaacaa agttatcgtn nttgaagaga anattttttt ggnttngncc      120
cgattaagcg ncaaatgtgt agcaaaaangc cgtgccactt gtggcgtagc tncgtcgggt      180

```

cgattcgacg	acaaggcgtn	gcgcgntanc	gttagtctcn	aatngaccn	gtggcatgag	240
cccacgangg	nttcgtgtcg	tcacatggnc	tctagacata	acgcncncn	ttttttncag	300
agggggntgc	cgccttagg	gaggnagggg	tggggacact	agccaancca	nantctnacc	360
ccattgaaga	aaagg					376

<210> 595

<211> 242

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(242)

<223> n = A,T,C or G

<400> 595

agnctgctgn	tcgtnccctn	tatgtggctt	catnntgagg	acaanagtng	cactgaggct	60
tgngnatgcc	aggcaaggnc	aagctggctc	aaaaagcatc	caccacctc	tgnaanggg	120
atgccangag	cangtgcacc	agtcccaact	angagnccn	ggcatgntac	atcttcttcc	180
accctnaaa	ntttgngcta	caangnccat	ttttcttttt	ctcttaagg	ncnctggct	240
tc						242

<210> 596

<211> 535

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(535)

<223> n = A,T,C or G

<400> 596

accagtggga	tactgctaaa	nagatatatta	tgcagcctca	tatgttaagt	cgtatatttt	60
gaaagctttt	taaatttttt	ctttaagaag	atttttagatg	cttatcactg	agtaccagag	120
ggatgtaggc	tgatgccctt	atcaacaaaag	tcagggactg	tggcacacaa	ggattgacta	180
ctgcagacac	ggccacaatg	ctacctctag	agggcctgaa	tccccctgcc	ctctctgggtg	240
gggagaaggg	ctggcagagc	cattagcatg	ggctccggcc	aatcctggcc	actttgacac	300
tcttggtgct	gacccagggg	cctggaggaa	gggatgaggt	gggcagtaga	gatgctcagg	360
gcagtggccc	ctttccatcc	acactggaac	tatttcagta	ttttaccacc	aattcagcca	420
ttcccttggtg	cgttggtgta	acatcagccc	tgctccaggt	ctcagtttcc	cctttgtaaa	480
gggaaaagctc	tggattcagg	gagtgatgaa	gaggtcatca	tggctcttgag	aattc	535

<210> 597

<211> 257

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(257)

<223> n = A,T,C or G

<400> 597

tttcnataacc	caaaantacc	ccatattang	accanacatt	tgtctnggaa	aaattaccat	60
tntntaacnt	ttgggccacc	tgagannaaa	tgggtgtaat	ncatgataag	atggancagn	120
attnctctta	agatnngatn	agaccccggt	tttcacggaa	catatccaag	nacccaatag	180
gnaacaagcc	acgggnggag	tcacaaacat	atattcttta	ctctcataat	ccgtmncaca	240
naactnttgn	acttgac					257

<210> 598

<211> 222

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(222)

<223> n = A,T,C or G

<400> 598

nntggntacc	gtcnaaactt	nncttggtac	ccgagctcgg	atccactagt	ccagtgtggt	60
ggaattccat	tgtgttgggc	tataagctgt	aatagtggag	ncgtgctngg	ttcattgcan	120
nagnccctcc	gcanncacnc	ttgnnacaac	ctgtgagnag	gcnataaatc	atccacataa	180
tcatactctg	atgaanctga	ctcaaacgca	tcacantaca	cc		222

<210> 599

<211> 238

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(238)

<223> n = A,T,C or G

<400> 599

gcatgacatc	ancgatgtnt	ttggnnacct	ganattngct	aaaactngng	natgccgggn	60
atgnaggttt	ggtantgatc	tatgcactca	catctcatgg	ggacgtttca	tgtggagtgn	120
tcgacaangt	tgctgnancn	gagaagtgat	gatctcagtt	gaaaggggtca	tgtgaataca	180
cnttacactt	gaaaaagaag	cacattggga	atatcacgaa	acgnccacca	acatcctg	238

<210> 600

<211> 232

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(232)

<223> n = A,T,C or G


```

cttctcttct taaaatngaa aaaaaaattg tttaaaccce naaggtctga ataccceagc 780
nccctgaach anagaacaan gccggagcac cccctcccaa atcccc 826

```

<210> 603

<211> 817

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1) ... (817)

<223> n = A,T,C or G

<400> 603

```

nnangacttt tgtggtntta tacaattntt ttttctattt ctatgaagag aaagccacag 60
agtcctaaaa taattctaaa actcatcatg actttcttgc ctaaaagatc ttgatttcaa 120
tcgtgcctag ttttgcttta atcacttgct tgagaaatac ataaatcccc acttaagatt 180
agtgacagga tatctctggc acccatttct ggttctatta aaattcctag agatgtcaaa 240
aattacatta ggccacctga caggctatac ctagaagaga aaaaatgatt tgtaaaagca 300
gtggggctat ttgcgattgc tttttttttt tottaaatat cacctattag gttgaaaacc 360
tgaaattgca gctttctgta gaaatggcgg aagacaaact aacattttta aagcgtctct 420
atthagctct gatgagtact acacccctga tattcttctg atactaaaat aattttccta 480
gtgtagtcta aactttttta aaaagacatg taatccgagg agtttgtaac tcaaaacgag 540
tgcatctagg aggtatcgca agccgtttct ggattaaatt ccagctagc ttgcttgctt 600
agcagggggc ggnaaanaag acatctgcag cctaggggaag aaaacctttc gcattgttct 660
tacgtgttta cgttatttta tttcctanaa caaggcngaa ttgggactcg aatgggtcag 720
ttgggggtgg ggatccctcg gtncataaaa ngtcanaaag anggtacagg cggaacncca 780
agggtcgtcc tgcatttana ctcggaattt tgggtgcc 817

```

<210> 604

<211> 694

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1) ... (694)

<223> n = A,T,C or G

<400> 604

```

cttttcaaat catttttnct cttctaggta tancctgtca ggtggcctaa tgtaattttt 60
gacatctcta ngaattttta tagaaccaga aatgggtgcc agagatatgc ctgcactaat 120
cttaagtggg gatztatgta tttctcaagc aagtgattaa agcaaaacta ggcacgattg 180
aatcaagat ctttttaggca anaaagtcac gatgagtttt agaattattt taggactctg 240
tggctttctc ttcatagaaa tagaaaaaaa aattgtataa aaccacaaaa ggtcctgaat 300
agccaaagca acactganca aaaagaacan agcaggggaag caacacacta ccngaattca 360
aattatacta ccaggggtgta gtaaccacaaa cagcattcta ttggcataaa atagacacca 420
agaccaatgg ancagaataa agaacccccc aaataaatcc atatatntac cgccanctga 480
ttatcaataa cnaacaccaa gaacatatnt taagggaant nctattcaat aantagtgtc 540
ggnaaaaaact gggaaatcca tatgcagaaa naatgaaact agacccttat ccctcaccat 600
acgcaaannt caacttcgga atgggattac aaaacttaag acattccaac ccaagaaact 660

```



```

<400> 626
gcaacaatca gatcatgtta aagtaaatct ccattgccct ggatcacttc aggatttaat      60
tgtccaagga gagcagggtt ctctgtgaa aaaaagggtg ggaaatgttt gagagtaaaa      120
aatacaaaat tcaaccggtc gaaaatacac cactccattc agtgctctac ccccataagc      180
c                                          181

```

```

<210> 627
<211> 813
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(813)
<223> n = A,T,C or G

```

```

<400> 627
accaagctgg agctcgcgcg cctgcaggtc gacactagtg gatccaaagt gaacgtgaag      60
gtgagcagag gagaacttgc gatggcaaag ttaaaaacaa gaggagatga tggctctgggt      120
gtggcacagg atgttaaaaa aattctcctg tccttaagga gttactgcta tttgagtaat      180
gtgccacttc cctacatagc cttctatgca gaaatgctat atttccactt cacaaccag      240
aacgtgcatt ttattttaca tttagaggag gaacaaacaa ccagaaggca aaaactgggtg      300
cattattttt tgcaattctc ttggaaagag ttcgttttta acttctgctc agacagcaca      360
caactactgg gaatatattt taatttcaaa tctgatgtgt gacatctggg aactcattta      420
ttgctaatag agttttcaca ggaagcagca gtcaccagta gctcatctta tttttcagtt      480
ggcaaagtgt tgtttacctt ttattggcct gcacgggtgt ctcttatcac aggatattta      540
attagaaaac gcaagtagcc taacatagaa nagaaatgga gtggtagata atagtagata      600
gaatggctaa atatttttat tacagtgatg taatatcact gnaatttatg gttaaaaatt      660
atgtaatact caaaaggaat tctcagactg gcgaaacagc tggncacagc ctntcacagg      720
gctttanact cctnttgagc tttccccctg ntggacttta gtcttccttt tacncccgna      780
gttnccattn nttaccaatt gtnccgggaa ana                                          813

```

```

<210> 628
<211> 646
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(646)
<223> n = A,T,C or G

```

```

<400> 628
tttggngnngn ggtgtctcnt ttgggtggac tttttgggtc gtagggcccc aaggccgtta      60
atcccgtaat aacggaagac gaagaagagt cagaagagtg cttctataag gatcgggacg      120
agactacctt agaggaataa aggaaaaaag cagaggagga agagtggtag aaggagtcag      180
aagaaacca cactgcgttc tgaacctgga gccttatcaa aaaggctctag ataaacgata      240
gcgatctcga tatcgagctc aagaggtagg tttagagact tctcgtcctc gagagcgaaa      300
tggaagatct cgacgacgat aagaagttaa agtgtagagg gtgcttgagg agcgctgga      360
aggattctgc ggaggggaccc atcgacgtag agacttgaag gcctactaag gtccacaaga      420

```

```

agccccggctc tttctccgaa tggtcggagc gtacagtatg cgacgtcgat cggcagacaa 480
gctggcggtta gactcgaagt gttcgggcga atcgacttat aatagtcgcg cgctagtaac 540
gtaggaacac gaagagtagt cgaaagaaaa cgttttagtga gggaaaagat tagggaaaaa 600
ggagaggcctt aataactaag acacttggag cctaggccaa cgcgaa 646

```

<210> 629

<211> 617

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(617)

<223> n = A,T,C or G

<400> 629

```

gccccncccc cctcctnngg gcttatnngg acagaccac gtagtactct aaatcttctc 60
ctacgccgga caacggaccc tataccaatt cgaatcttgg aactccgac cgccggattc 120
tcttccccctt tcggcttccc ctttctgtcg gtacccctcc ctagtctct cctacacctt 180
cgtaccgtcg atatatagtc gccgcggact agcctattta ggtgtcctag actcgttatt 240
gatccactca ttagtctagt actatgcgtc acgtatctta gttgcctaag agggagatta 300
aatcctccac aagttccgac gaattcctgg actctcgtac tagcaaactt tcttatgagg 360
cttccttgta tatcttctgg atgtttctcg tgtcccggtc ctccgctact actagagctc 420
cttgccctat ctctagaagt agaggactct cgggttcggt ctccaaatct agcgctagag 480
ctatcgctac ccgctcgatt cccccagcgg aatcttgaaa cctgaggtag tacacaaacc 540
ctcncatct tcctcgggtt gctccttctt ctcatcccc cttcccgctt tctcgggaan 600
gaatctactt tancttc 617

```

<210> 630

<211> 644

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(644)

<223> n = A,T,C or G

<400> 630

```

cnntcggcnt gggttttntt ctgagnnncc ccccccccc cccccccaaa cttacaccca 60
ccaaacactt tccgccccct acctaggaga cattagaagg gtttaggctt cggcgatatag 120
taaagtcctc tacctcggaa gtagagaatt cggtatTTaa attcagggtt agaggctcgc 180
tcgttagatt tatagtttag gtttagaata ggaaaccttc gatcttcctt agaagggtaa 240
taagttaggc cctaaatccg tctaaccaag gcgttaaggc ccgtacctaa acctagtctt 300
atcttctatc aggcgcacca atataggttag gttctacttt cgtataggcc ttaaggaata 360
gttcggtagt tatcgaaggc actcctctct aggctaggct tttctcagtc ttagtactcc 420
gggaccgtcg tcgcanaaat atcgatggac ggtagggtat tccgcgttac gcgtcgggct 480
agggatatag agcgaattat cggcgagagg cggtcgctan gaatcgggat caatatgntg 540
ttctttaccc tacggatatc ggcagaaaac ataaaacctt ctnaccangg ataagggtat 600
atcggacccc taaaataaca gtaacattta gantactagt accc 644

```

<210> 631
 <211> 526
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(526)
 <223> n = A,T,C or G

<400> 631
 ccntcggcctt ggggttttttt ctgagccccc cccccccccc cccccccccc cccccccggc 60
 cccatagccc caccggnccc acccaaattt taacaaaata aatntaccta tcgntcacct 120
 atcccnctga tcgngtaggt cggtagccgt accgngatc ncnacgattn ttcgggtcgt 180
 cncccttaan acggncccggt agccnccgga anaaatacta cgagngactc taatntagca 240
 anaccgccc tcnattanta gcaccccttag tcttccaatg ncnnggattn ngaatccttn 300
 naagttatcg ggtagaacgg gtcccgggtcc cccgccctct ttncaattaa cgcgggttac 360
 aaantcgggt tctaaattcc ncacgaattt ngncggcaac attcncgggn ccttattanc 420
 cntttccaac cccgatacnc nagctcgatc gggcctttanc gaatccgggg tcnccccga 480
 ngantccggg tcctttgagt ngctctagga cgggttacgac ggagga 526

<210> 632
 <211> 647
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(647)
 <223> n = A,T,C or G

<400> 632
 tttggngggc gggngctcat ttgggtggac tttttgggtc gtaggaacct ggtatgaggg 60
 gtgttttgag tttcttcttc gtctctctct ggaggttcgg tttcgattga gattcgggtt 120
 cgtctttatc ttacgaggca ccctgatatt gttgcctttt ggtttggttg tggagagttt 180
 tgtctactc tagcgggtca tgccgatgat atgtagcctc cgtggcctga tagtgatgtt 240
 gtgagcttga gaggggagtt gtgggtggtt cgggcggagt aggaggggtt ggagcaccgg 300
 gattgggaga tatagaatca taagtgttag gtataggctc attgagcgag ttcgtggaat 360
 tcgtgtggtc atcataatta gagtgaggat gggctctata tttcttagag gacgcacggt 420
 cgtgattcgg ggtttgatgg gtgttcttct tgtgggcacg attagcttgt tcatgatggt 480
 aaggaccata ctgtttcgaa tgaggattcg tgtcttcgga ttgttggtga tattgtggnc 540
 tanactatct agtgtaagcc ggaggtggtt tgccgtgggt gagtatccga nnttcattcg 600
 ganggtatgc gtgcggagcg gtcctttagt acattccgga aaaatgg 647

<210> 633
 <211> 630
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature


```

agatacccaa agaatagttc cactcaactt cgtctaagta aaactctaga acttccaaac 120
ataaaagact tcgcgcggtt agctacacag cctacgggaa tctcacgaat cccgattcaa 180
gtcccactct cgaccacacc ccggtatcgt cgttttccca taccaatgtc gaaaaataaa 240
ataaaatcca gtcaagcccc acggtaaagc ggggtagggc taggcgaaga ggcaggaacc 300
gttcgaggcc gggggctttc aaaatacaaa acaactactt aaagtttacc ctttctaaag 360
tcgggggcaa cggttaaagc acgcctctaa agtactactc gtttcgagaa ggggtagtca 420
tctcccgcat agagactctc gcgtatatca actcgcacgc cttctagcat tccgacggtc 480
gcccgcggtc acatatcttg cggattagct ccgagggact atagggttaa ttagtctagt 540
aaattctctt agaggatagt cggggtcgta gttaggcagt acgaggggac atggnctgcg 600
tcgtgctcta ccttgacagc atactcttat aaacatcttt ttctt 645

```

<210> 636

<211> 643

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(643)

<223> n = A,T,C or G

<400> 636

```

cttcgggtt ggggtttttt ctgaccccc ccccccccc cctagcggaa aacaatcccc 60
accgagattt tattaatcgt aaaactcgcc ttoggtacca agtcttctc cttcccgtaa 120
cctggctccc tcctagnngc tttacgaacg tccctctctt tcttacggct cggaagtggg 180
tacggttaaa tccggaggng gggctaacga atccaaggct aactcctctt anagtttggt 240
gtccnncnct ttagtaagga tccgtggagg gcgagtattt gncccccggc ctttattnta 300
tagttcccta gtacgataaa gntaccggct atcctattac agcggataaa agttatttan 360
agggccgacg tcncgcgtag acaggctaca gctagnngag gtaccgcctc cgactantcc 420
gttgnttccg acaagggnagt ttcggttaac tccacaaact cctccgccga ctctanggtg 480
gggacggcag ttccnncggt tagtggtgct tatagagaag ggcatttgag ttggacgtta 540
cnttttaaca taggttattc cgttttaggt cttgcgggcc cgtgggggta gtnncceggc 600
gcgttnntat cggcgatttt ccgcagtttc cgtttccggn tnt 643

```

<210> 637

<211> 631

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(631)

<223> n = A,T,C or G

<400> 637

```

gggttntctc atttgggtgg actttttggg tcgtaggaac cggtatgnag gagtaggagt 60
cgctgggaag actagaagtt agctacggac gattagtgtg attccactct taataacgag 120
taatcgttta cgtcgggttg gtgtttcggg gttttggaga gtaagcgtag ttgtggagtt 180
tcgcatatag gtccccttac ttccggcgatc tcgtcttctg tcggttaggt tattattggt 240
catccttcgc attagtagta gggttggctg gataaatcga tagctattct ttagaattcg 300
tagtcggaga attcgtgtac gaagtccttt aagttcttta agttcgcgag taagacgtgt 360

```



```

acggttatatt tgcgctcgac gtaggtgtcg ttacggggag tttcgtttta ggggttttacg 420
tagaacgtta ttaagcacgg taatacgata gaggattacg cgacgtattc gtcttagaac 480
gtcgattttt cgaaggcgca tttgttatcg aaggggagtc cttggagaat cgagatattc 540
caagaatatt acggagatta cagatcggaa ggctcccagag atcggacgta ttaccggtct 600
cgccccgaaac gagtaggtat cntccggata a 631

```

<210> 638

<211> 606

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(606)

<223> n = A,T,C or G

<400> 638

```

ccccccccc ctcaaccatc nattccccac ctcaacgcga attacggttt cgaaagtcga 60
caataagtcc ggtcgagtag aggggaatcag gggctgggtan aaaggaccac gggcggaana 120
taccggtctc cttccgggga ggcacgtcgg ggaaagggaa gagagcggtc tagttcgtag 180
gcaaacaggt cagaaaagt aagggttaaag gtcggagggg agaggatagc tagtacgctt 240
agttcggggc tcgggcgcag ggccactttc ctcttttcgcg ttcctttact ctgcttacga 300
gttcaggctc cggagttccg cgccggaggt cgtcgcgacg ctaggaatgg ggactcgcgc 360
agtccccggt tacccttcgg gattctatgt tttcgcgat agacggagac cgggtagtag 420
ggttccgctc taccgccact cgtcgccttg atccggccc ctcgccttaa gggcgatgaa 480
agattaggtt ttagggtctc acgggacgag gcatagggcg ggagaagggg ggaggggtcg 540
ggggtcgaag ggantaagaa atcgcantcg cgcgggggtcg gtagganccg aaatttttct 600
cnnctg 606

```

<210> 639

<211> 592

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(592)

<223> n = A,T,C or G

<400> 639

```

tcctcgggt tgggtttttt tctgagcccc ccccccccc cccccgggaa cgagaaaaca 60
atcccaccct accgcgggga gtgggttgna cgcttagttc tagaatcctc ggaatcgtcc 120
tcggcggttg gtagttccgg cgattccgag tatgccgaag tgtatcgctc cgtctagagg 180
ttggtatctg tttatcgca tgacgtatt gactcggatg ctttcgaagt agggggatag 240
gcgcatagat acgcctccgc ggtgtcctct gaagtggccg catccgtgga cgcagcgtag 300
acagctctgg tggacgataa cggtctctcg tactcctact ccggctatta tgtagagag 360
gacttgtttc tgaacggata taccattagc gaaggggtac cctccgctaa cgcaggcgtt 420
tctaacagtt cttccgggcg ctccgaattt agattgacgc ctccgcagca ttgtgggatc 480
ctcttcggtt agccctcttt ataggatttc tcctccgccc cgaaagangg ctggtcgtcc 540
ccggcangta tgtctagctc gaacgctttg ttactccttt gttttcgaaa na 592

```

<210> 640
 <211> 637
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(637)
 <223> n = A,T,C or G

<400> 640
 ctttgtggcg gtggtgtgtct catttgggtg gacttttttg gtcgtaggct tatccgggtn 60
 gggctcccca agtagcttag gatcgccggc tagttccggt cccgcccgtc gaaagcgcg 120
 ttcggcgggc ggccccggt tegtccggc gctttaccct catagagtgc caggtctcgg 180
 ttcttacggg ttctgcggcg atagatttta cggcgagagg tcggtatctt cgccgcttta 240
 cgttcggtcg gcatctacgc ctagttcaca ggtagtttat gcgccggagc gcgtgacgga 300
 gaggttatac gggacgcgga agaaccgcct ccaaagtact agtacaggct cgttcgggcg 360
 tagatctcct cgctcggtcg gcggttctta cttctagggc cgctctacgg tttaaggcgg 420
 tcgttagatc tttagaaacta tactcaagtt tcagtcggaa gaaaggaagt agagagaagg 480
 gtaaacgatt acctccggtt ctagcccttt ttactcgcat aacgggagaa cggggtccgg 540
 ctctcagata cgctcgcga gacgtcgcga ttcaacttta acctccgcta gggcatccgt 600
 atacggttaa cgcggtaaaa gcgacctcgg aaacctc 637

<210> 641
 <211> 649
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(649)
 <223> n = A,T,C or G

<400> 641
 ctntgtggcg gtggttgtct cagtttgggt ggattttttg gtcgtaggna acctggtatg 60
 aggtctagtt tcttcaacga ttcttggttc agttacgcga ccctatcctt atcttacaat 120
 gtcttctaca tcaggttcat caattaatat atcaattaca cattaacgac ggtgtgacgc 180
 aatatgagaa agtatacatt aagggtatta tatattatct gcttaaaaag gttcctgaca 240
 tgggacaact tcacccacca ttctagaagc ccccccctct gtaggacccc ctcgagttcc 300
 ccattatctt agttcagttt tcattttttta accaggaggg tatcggtttt taataggtac 360
 tattttgtca aacttttctag aagcttttat ttcaaataata cttgcaccat ctgtactagg 420
 agcactaaact attcgagtct attacagctc aacagaaaat aattgaaatt aaacaaccta 480
 agtatcgctc accataaccc catcgggctc tcaccccat tcttcataag ttctagagca 540
 tcttgagctc ttctctatta ccttgatgg tactcatggt ctaatacccc ccgcagttat 600
 aggtccttat ggatcctatg ctaccaaccg tctaatecct tctatcacn 649

<210> 642
 <211> 645
 <212> DNA
 <213> Homo sapien

```

<220>
<221> misc_feature
<222> (1)...(645)
<223> n = A,T,C or G

<400> 642
tccttcggct tgggtttttt ttcgtcgcgg gttactatta tcgattgtta cttgtaaagg      60
cgatactccc accgctcacg atattagacc tgctcctcta gaagcgaacg gcgataggctc      120
tactcggccg gcgaagacgg cgaacgggta ggaggagcca tatgcaaccc taacgggagat      180
tataagtact gggaaaaata ctagtattaa ggtagcgggt taagataggt ggagagacac      240
tattcacgag cataagcact tagaagggtc tctcgaggag aggtaggcta cggactacgt      300
tccttccttc tctagcctcg agagggagta tagatgattc gcaaaagaga atccctccta      360
tacgctggca taactagacg acgcgtcgtc gggaaatctc gccaaccccta ttgcgacctc      420
caaaaggaag attgtcgttt catagaacgc taatactcgg ggtcttcccg aatcatagcc      480
gcatatcggg aagaagacgg taaaatcgcg cgattctaac aagattctgt agacttaagg      540
ctaagcacta gaagcgatct cgattccgga tcttaagatc atactaatag ttcggtcaca      600
ccagacgacg attagccact agaagcccta ctccgtngaa accggg                        645

<210> 643
<211> 586
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(586)
<223> n = A,T,C or G

<400> 643
ctttgtggcg gcggtgtctc atttgggtgg atttttgggt cgtaggaacc tggatatgcag      60
ggtcgcgccc gaattaaaaa cgggatcccc aaaacgnngn ttcgcaagaa gagaagaatc      120
atagcgatag anctttcata gtacaaagggt aactaagagg aaaataatgc agattcagaa      180
ctagttgcc aattagaact cgattaggcc aaggatccga gcctggcgct atcacttcgg      240
gacttaagct acggtagagc agtcggtcct gaagcatagc tcccgtagga cgtaggaaac      300
tagtcgggca cggaggacat actctcgagt ctcggaacgt ctatttagaa tataaacgca      360
ttaacctcag aaggcgccga cgcggttact ctctagggaa ctatttcatt ccttcgggag      420
ctcccctatt tttccaacac atataccggc aaaggaaaaa cttntgtcct cgggtctaaag      480
agagggaaaa aaaacgatat ctaggttcgg gtttatccat ttaaaaaanat ngacgcgact      540
actccctttc aaaggggagtt tccccctagg nagagttcaa cngaag                        586

<210> 644
<211> 646
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(646)
<223> n = A,T,C or G

<400> 644

```

```

ctttgtggcg gtggttgtct catttggtg gcatttttgg gtcgtaggaa cctgggatng      60
agggtcattt gacttgtttc tcaaatccca tggatatggtg ggtggcgtgc ggggtggcgg      120
tcggttcggc ggggtggggg gtcgtcctcc aaaggagttg ctagagggct tttagtggtt      180
ttagggcggg aaggggttag agcggagaga cgtcgtcgtg gaagcttctg gcggagcgcg      240
agaaggtagt tagcgcgggt tcggaagatt ctcagaattc gagaagaggt agtggggcgc      300
ggagagagag tttctaagtc taaacgtaga ggtcgtccta gtcgggcccgg gagtagcttt      360
taagctagag gtcgaggtcc tcgtttaggc tccgggctct tcgggcagta tcctctttct      420
cgaggaacgg agcgaccgac gtcgtagccg gaccgctcta tccgtacgtt tagagatacg      480
ctcacctcca cgggcgtata tgcccgata cgtataaacg cgtaatatatac tcgcgcgtaa      540
aacacgtata cactatatac acgcacgtata cggaccgtat agcgttatatac gcgcgcgtat      600
attaatttac acttatatac gcgttaacac gatatatcac acnccg                      646

```

<210> 645

<211> 654

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1) ... (654)

<223> n = A,T,C or G

<400> 645

```

ncntcggct tgggtttttt tctgaccccc cccccccccc cccccggctg acaacgtgcc      60
caccgttgcc atcccagcat agctgggtcg ttctgtttta ttcttagtag tttagttcgc      120
ctatagtcct tcgtctatcg tctatcattt aaggaggcgg ggctcgctct ttagggcggg      180
tatcttaggt attcttctgg ttccggctgc cgtctcggag tctggtcctt ttgctttcct      240
ttcttggtcg aacttcgtgt ttgatcgcgt tgtttctttg gggtcgtcat acctaagggc      300
cacttcgcca acaaacaggt ttgtgtagtc gtttctatta gggttcgtg gccggcgctc      360
ttactggttg gcgattttta acgcgttttg ttttaatttg cttcctcccc tagggctcgc      420
tcggctcttct ctctgttcgc tgctctcgtc cggccttttg tgcggggata gctccggcta      480
ttancgtgcc gtgtccgtgt ggnttttgtc caatgtgaag gcctaggggt gcgggcttct      540
ttggccatgg nttccctctt tgtgancctt aggggtaacg antcgttaatt naaggctcggg      600
ggttggnata cgttntangg gangcctgng tccgntatct cttgttttgg cctn                      654

```

<210> 646

<211> 645

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1) ... (645)

<223> n = A,T,C or G

<400> 646

```

tccttcggct tgggtttttt tctgagcccc cccccccccc cccccacgcc aagtacacag      60
acccacccaa aacaacgtca acacaacttc ggggtatacgg accttaagag agaccccgtc      120
gtagacccta ccacagccat ccaatagtc aacaacaagg gcgcacccaa tccatccata      180
gagctatcaa acaacggagg ggaaaggaaa gagcagggtc aacttagcag agatcgaagt      240
cggcactaat tcctttcaag tactcgctcg gctttagatt cggggtaaaag tccgctctca      300

```

```

aagggccaac gaggttttaa agcgaccccc gtatcgagtc ttcttcgtat tcattaaggc 360
gttaaaggta cgagacctag aagagagtag aattagccca ccaaatcgcc taaaccggca 420
aaaacgacca aaagtcaaag acccttaca atatacactt aaaacgcca ccccaaaaac 480
gcgatcagta acgcacgtac ctttcccacg cttttctttc tttcactctc caaaacaaac 540
ccgaatatatt agcgcaaaaa atatacgagg gagaattaga agctattacc cgaaaaaaaa 600
ncgganangg antaaatngt ggggaatana cgtttggttt ttctg 645

```

<210> 647

<211> 753

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(753)

<223> n = A,T,C or G

<400> 647

```

accttacctg gtaccgggcc cccctcgag ttttttttt tccaaataca actcagattg 60
tatacgaaaa gctgataata cattgacttt tgctgtttta atcccttgag ctttgataa 120
tgattttttt tgtgttaaca attgtagtat ataaaatcgg attcaccatc cttctgatgc 180
catattgatt agtttgattt tatggtgatg ggatcattgt gtgttaactg tattaagaag 240
aaatggattt gattgacttt gcatocattt ttatctgtgt tactttcatg tttatattaa 300
aagcatttct ggaccagaat aagttaagt gtataatttg ctttttacac gtttatataa 360
ttgaagttag caatgtggca aaatctctaa tggaaataaa atgcttcaga atgatgacat 420
aaatctgagc tatttcttgc ctggagaaca agtggttattc ataataattt aatagcttct 480
gaggtgtttt gttcatgtga tgaaggctta tccaccttgt atcaattcat gggctctgct 540
ttgtttaatg tagtcagggt gttaatacna gacttaagag tcatacctact gtgataagt 600
gtgagtgaag attacatgtc ttangaaaat tatactggga atactcttga cattaatggg 660
tttaaagtgt ttaaggctag gggatgatgc aatgganaan atncttccaa angtttctgg 720
ttgtttatat ttgnngaagn catnaagana ccg 753

```

<210> 648

<211> 383

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(383)

<223> n = A,T,C or G

<400> 648

```

gatatcccg ggaaatgcgg aggcctttng gcttacgtgt ttaccgcgta gggcaaagcc 60
ttgncaaatt cccggccagc ggagcggcga ggggtggggac tcacgggaag ttaaacagcc 120
tcgtcggcgt cctcgaggct ccaaaaccag gctctaggcg gggacgactg cagccgttat 180
ggaggccacc gcggctacgg ccgcggctga ggcctcccca ggtggagcgg tggcctggag 240
gggaatcttg atcctgggcc agccacctgt caagaggagg cggagcgtca tgcctctgga 300
agactggatg aatattctcc aggagcctga cgaaggcgaa gaagtctttg cagaggaaat 360
tgaatgctgt ctgatgctac aat 383

```

<211> 349

<213> Homo sapien

<221> misc feature

<223> n = A, T, C or G

cgattgtnta cnaqtcttag agtaagctta agntcgttac cgaqctcgga tccactagtc 60

cagtgtggtg ggaattccat tgtgttgggt cactagtaaa tggatttagc taqacanaag 120

anatttacc tttccattt agcacagtga gganaggcta nacagctagg atgcaataaa 180

aaaaatttta atgagaaatg tgtgtggtag attaatctta ttaatctcaa gttatagatt 240

aaaaaattta agtacncat aaatgccatt tgcctttgct aangntacat ttttatqaan 300

aangaccntg catacnaat ganatactgg actttnggna cttgangga 349

<211> 306

<213> Homo sapien

<221> misc feature

<223> n = A, T, C or G

cattgtgttg ggagcatcct tccatcagct cccatgagaa attctctgtt gggttttaagc 60

aatccccaaa tatatcatat tgacatgaat atatcatctc ctcaatgtcc agcattagca 120

gacaagatga gtgctgaaga tgatataact cctacctctt atgtaggcta gaggtaaaagt 180

ctggctctgc tgactgtggg gacataccga aaaggaatgt gggttaatat cagangacct 240

ccctgcagat ccganantca gggnctggac tttctgggan aggaagcnaa aagttatntc 300

tgaacc 306

<211> 769

<213> Homo sapien

<221> misc feature

<223> n = A, T, C or G

cattgtgttg ggcaggggtca tttctaaggc atgggctgga agcttttatt taaaacttta 60

catgtcttag aagcactctg gttgttgcta ggcagacaat tttacatctc ttgctataacc 120

agttgcatga agttcatcat gcatattggc tgtggaaaac cttaacagca tcatgtcata 180

aggtttcagt aaggtttaaa tgaaatcatg tattaagcac ttagtatagt gcaccttaaa 240

```
<210> 652
<211> 267
<212> DNA
<213> Homo sapien
```

<400>	652						
nnangccctt	taaccattgn	ggcctccacg	cnntggcggc	cgctctacaa	ctagnnggatc		60
cgcnactcta	gnanaangat	tggctcttnt	gggntgggcc	ggnccgggctg	gggcgttaag		120
cggggctggg	cgcgcgcgcn	ggttgnacna	ggcgccgcgg	ccncacacn	cccggagcac		180
cctcnttgcn	gcctntcccc	gctcaccccg	cgcgcgcgcn	tccgcttttt	ccncacccan		240
agcncntntt	atctntgtct	cctccgg					267

```
<220>
<221> misc_feature
<222> (1) ... (501)
<223> n = A, T, C or G
```

<210>	654
<211>	710
<212>	DNA

$\langle 220 \rangle$ $\langle 222 \rangle \quad (1) \dots (710)$

<400> 654

<210> 655

<212> DNA

<220>

 $\langle 222 \rangle \quad (1) \dots (202)$

<400> 655

<210> 656

<212> DNA

<220>

 $\langle 222 \rangle \quad (1) \dots (308)$

<400> 656

gctgntgaaa	gaccacaccg	aaaaactctn	ctttccgact	tccacatgat	gatngcatg	60
tgggtggtgag	agacttatca	tgacgacatc	gcttccnacc	atcgcanccn	ctgcccaagc	120
ccattcatgg	aggcctgggn	antttctgtg	ntgaentnga	cnctanacnc	tnccactqtn	180

tgctatccag	acttgnttng	aatatnttat	tggcnaaana	canttnccga	atgctgtgnt	240
tgnn cattga	angatctgat	cactatgaga	gggtgaggac	nncctgctng	ctggcantnt	300
ntaaccn						308

<210> 657

<211> 696

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(696)

<223> n = A,T,C or G

<400> 657

accntttcca	caatnctggn	ctccccgogg	tggcgggccgc	gtcgaccagc	aacctcagct	60
gtgggtcttg	ttacagtaat	gagttactgt	aaggaaagtg	tgacatttcg	agcaatttga	120
tttgtttaaa	aactagagca	gtttcagggg	tttccttgta	aatctgtctt	atgtgtcttc	180
aatgttcttt	cttgaggagt	agagaaagga	attgttagga	atgatgcata	aaccatggct	240
tattttatct	cgctgccacc	cataatcaga	gcagattctt	gggactatga	ccctcatgga	300
gacatgacaa	ttgtgtgtgt	ggtgggtggg	agaaaagagc	tggaattttt	tagggtctag	360
agggccaat	caggactatt	ttatggagct	ctgctcacca	actttaagtg	agcaccaggg	420
gtgngaaagc	gaatcttggg	ntcaaaanaa	caatggnaag	gggtaagttg	gtatnctgaa	480
ctggccactt	cggactctta	tttaactggg	tattctcant	taaggaggcn	nggggtggtc	540
tggcttgtna	aggaaaagcct	gtgcaatgga	atgactttta	aaccccccat	taaaaaaaaa	600
angntataaa	tcttggtgtc	taanaangaa	gcctgggttc	tnttanccca	ttttnccccc	660
gggaaggnaa	atnttcttag	gnaanggaag	ggaagg			696

<210> 658

<211> 698

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(698)

<223> n = A,T,C or G

<400> 658

ctggactccc	cgcggtggcg	gccgtctctag	aactagtgga	tccgtgttgg	ctcaattctc	60
aaggtctgtg	ctgtgcggcc	tgttccccac	acgtgctgct	cagctcaggc	aagcaccgag	120
cttgtgttgt	ttcatgctca	gcgtggaggc	ccctcctcca	ggtcgctgct	ctgtgggggt	180
cccatacact	caggctccta	ggaggagtcc	atttagaaag	ccagggtttt	tctcagagtc	240
ttagttcctt	gtgctgtcat	ccatttcaca	cgacttgggc	cctgctcggg	gcaacacagc	300
aagagaaaag	acagggaaaa	taagagaggg	accttgacac	cacacgctct	ggaccacaga	360
gccctgtgcc	cagctcctct	gtcaatacac	gtggaatctc	gtgcaggatc	gcaggggtct	420
gtgatgccac	caaagagcag	gccgggacag	ggttaggaga	gaaaggagag	ggaagtgggg	480
gtttctccta	cgcactctta	tttgagaggg	gaaaggcggg	tttgtattgg	ggttgtcggt	540
ctttgcaccc	acngcacagt	tgtgagacac	ccccatcctn	agatcaaagc	cccacataca	600
gcttggggaa	aaacaaaacn	aaacaaaaca	aaaacagtaa	acctccatgc	canttggttg	660
gnaagttttn	aatttncctt	cccnacccan	cttgccttc			698

<210> 659
 <211> 750
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(750)
 <223> n = A,T,C or G

<400> 659

ncaanctggn	ctccaccgcg	gtggcgggcg	ctctagacta	gtggatcctc	ctcatgggcc	60
tggatatctc	tgaacatatg	atgaacattg	cttatgaaaa	attattttgta	ngaaaattgt	120
gaggcctaag	aatgntattt	tcttttagtg	atgggtctttg	tttgcttctg	taaggnaactt	180
gtgggcactc	gtaagcttgg	atctctttta	tctaatacca	gntttgagat	tttcttggcc	240
ccatagatga	attaaaactg	gcgtacttct	tgtttacaag	anggataagt	ctcctagggg	300
aagtcttttg	gggtcccaag	tcaaaaagat	gagggattta	ccagtctctc	aaccttggtta	360
gccccagact	ccaaactttg	ccttctagtc	ccaagaggct	atcaaaaagc	aaaggccatc	420
ttccaccttc	ttttccanaa	cagcacacat	tccagacagt	acttgaaagc	aggaacctcc	480
ttatccctta	aaaacctctt	ggaancatct	tccctctctt	gcttctacta	tgcttggccc	540
acctancatt	cncntttttc	tggaaaccgg	aaaaancttn	tgacttnngt	tggctacatt	600
cagcttggcc	ccctacaatn	tggtttccat	ctgccctaan	gaaattttta	agggcacttt	660
ttttntggcc	cctgactttc	nntttttagg	gctttccccc	angctttgcc	cctttgggta	720
aaggggttat	tttcttcccc	cttttggaag				750

<210> 660
 <211> 849
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(849)
 <223> n = A,T,C or G

<400> 660

tcggatccac	tagtccagtg	tgggtggaatt	cgcgggccgc	gtcgacgggc	agtagtggtta	60
tgcntntcta	aatgttataa	ttatttcaga	attactctgc	cagaaagtta	tgatcataca	120
tagaagagtt	tgtagctaac	tttgaaagta	gtggaaagtg	gttttcatgt	attgtttggg	180
ttaatttaat	tttgattata	tttgggtttt	agttcaggta	atttttttgt	tgaaaacttc	240
aatgacaat	ttcttcatgg	ttactaaaga	tactcatgt	ggagtagttt	cagatttttt	300
tctgaataca	tgtattactt	ttagagatgt	aaagatgtga	aattactaag	agagaaaccc	360
atgtgatttg	tttagtggtg	caaaagtcgg	tagctccttt	gacctaagt	gccactgata	420
gttaaataga	tactgaagct	atgggcaggc	tggattgata	agaaaaagg	agacagagaa	480
atgggaaatt	gggaaagaac	tgtgcaaata	ggaaaaggag	agagcaacag	aacagaatta	540
gtaccacagt	gccgaagtgc	cacctcaggt	acttccatct	cccctctcct	gaagaattca	600
gtaacagttt	gcaaatggtc	aacacaatca	tttagtgatc	ctggttgata	ttttcaatac	660
tttctgggga	tttcttggtc	ggnttcaaaa	gatgatgctg	atagttttat	tgccccgtgaa	720
ggtattctga	agnttancat	aattttattg	tcagtaaaat	atttgaataa	aagngganga	780
agggaaatct	ggcntcttat	tttgggatnt	cngcnggggg	aangaggata	taattnaccc	840

849

<211> 653

<212> DNA

<213> Homo sapien

<220>

<221> misc feature

 $\langle 222 \rangle \quad (1) \dots (653)$

<223> n = A, T, C or G

<400> 661

aacttaagct	tggtagcgag	ctcgatccc	tagtccagtg	tgggtgaatt	cgcggccgcg	60
tcgacctcca	ttcgttttctt	gtcctttttt	ttcatttttt	ctcatgttct	attcacttta	120
ggtttctaag	ataaatatta	taaaataatt	tttacttata	aattattcac	tgataccctg	180
tctttaacat	gtgaaatgaa	ttcaaaagga	atcttaatga	gaaataatat	actcatgatg	240
tttaatagat	ttgatttcga	aataataagc	cctctgaagt	cctaagttaa	aaataaagca	300
acttgtttga	taattttttca	tcaagaatgt	atctgagtct	ctgagtaatt	attagtagga	360
atattccatt	atcacaaatta	cacagtataa	gctatttagt	ctaactttac	caaaaaaggg	420
agctacttta	acactgtgtg	agacttttta	tgggtttgca	ttgggtatgc	actattagca	480
agataaccta	ttttacagca	gtggttntta	acctttccca	tttatttgaa	aggcagctaa	540
gatatagtag	ttaatntaan	gggctgatgc	atttatatta	catgtagana	atggggagata	600
cnaaaqqqag	nqqqqqqana	ntttttqnat	tcnnaagcct	cnttqncaat	taa	653

<210> 662

<211> 646

<212> DNA

<213> Homo sapien

<220>

<221> misc feature

 $\langle 222 \rangle \quad (1) \dots (646)$

<223> n = A, T, C or G

<400> 662

aaacttaagc	ttggtaccgc	agctcggatc	cctagtcacg	tgtggtggaa	ttcgcggcgc	60
cgtcgaccca	gggacaggca	gccagnctg	gggtcaccag	ggtccctct	tgggccctcc	120
aanagcaaca	gtactggcaa	cagctgggat	ttgctgagca	cagactctgc	agcaggctcg	180
gttgagctct	ctgtgcctgt	tccttcatac	catcctcacg	ccatccatg	agatgggtcc	240
agctgttttc	agatgagaaa	atggcacagg	aagctggtaa	gtgacagtca	gaaatgaatg	300
ctggcagctt	antccttgga	cccaccgcag	tgcaggacct	tgtcacaacag	ggatcacctt	360
tgtccgccac	ctgttcatga	ggccacccag	ggtttgtgtg	gtcatttgtc	tcctttcctc	420
tgcttgctct	caaccagctg	ggctcattag	gctggggaac	ccagacccca	cacagtccct	480
ctcccgagang	ccagacacan	nctnccgcac	agnaaggact	tcagtccccg	aancaaatgt	540
ncctgggcgt	anaaactgna	gggnccccaa	tccttggtgg	ggtactgctt	tgcactggng	600
gaattcaccc	ctcattgnna	acctttccct	ntnnccacc	ctaaac		646

<210> 663

<211> 650

<212> DNA

$\langle 220 \rangle$ $\langle 222 \rangle \quad (1) \dots (650)$

<400> 663

aacttaagct	tggtaaccga	gctcggatcc	ctagtcacgt	gtggtggaat	tgcgggcgc	60
gtcgacgtcg	acgcggcgng	ccgtttcgac	gcagttgata	catattatta	tatactacat	120
nggttttcta	gaattaaaaa	attaatgtgt	agtgccagcc	ctagatgtaa	gttacatata	180
tcaactctat	ccaattttgt	cagccataaa	acttaccttt	ttcacatact	tctaactcta	240
acaatgtgag	aaatgtagat	cattgcaatt	ataccacaa	ggcagatggc	tacatgcaga	300
atggatagca	gaatctagct	acttacgcta	gccacatggt	agacgttttt	tcctttgttt	360
ttgcaaaatt	gcaatataag	ttgcatatcg	ttagagtga	aagatgtaaa	gaacccatag	420
aagccagtga	tgaaggacat	ttatatcttc	acctttacaa	angaccttaa	aattgcctat	480
gtggagcaga	aactggagga	gggcnaancc	atcngtaaaa	aaaattttgn	tnctatttgg	540
atttgggcac	cattattacc	tcccaggtn	cttttttgtn	ttaacctttc	ttttaaaaaa	600
aataattcnt	aattttttqq	caaaaaaaa	caaggttttt	atttaaaatt		650

<211> 678

<213> Homo sapien

<220>

 $\langle 222 \rangle \quad (1) \dots (678)$

<223> n = A, T, C or G

<400> 664

taaaaaatcta	gactacacta	ggaaattatt	ttantatcag	aagaatatca	ggggtgtagt	60
actcatcana	gctaaatgag	agcgctttta	aaatgttagt	ttgtcttccg	ccatttctac	120
agaaagctgc	aatttcaggt	tttcaaccta	ataggtgata	tttaagaaaa	aaaaaaagca	180
atcgcaaata	gccccactgc	ttttacaaat	cattttttct	cttctaggta	tagcctgtca	240
ggtggcctaa	tgtaatTTTT	gacatctcta	ggaattttta	tagaaccaga	aatgggtgcc	300
agagatatgc	ctgcactaat	cttaagtggg	gatttatgta	tttctcaagc	aagtgattaa	360
agcaaaacta	ggcacgattg	aaatcaanat	cttttaggca	agaaagtcat	gatgagtttt	420
anaattattt	taggactctg	tggcttttct	ttcatagaaa	tagaaaaaaa	aaattgtata	480
aaaaccacaa	aaggctctga	atagcccaaa	gcaacactga	acaaaangaa	caaagcagga	540
agcaacacac	taccggaatt	caattatact	accaaggtgt	antaaccaa	acagcattct	600
attgggcata	aatcagacca	aagaccagtg	ggaaacagaa	taaagaancc	caaaataaat	660
cctatatTTA	cnqccnc					678

<211> 694

<213> Homo sapien

<220>

<221> misc feature

<222> (1)...(694)

<223> n = A,T,C or G

<400> 665

```
cttttcaaatt catttttnct cttctaggta tancctgtca ggtggcctaa tgtaattttt    60
gacatctcta ngaatttttaa tagaaccaga aatgggtgcc agagatatgc ctgcactaat    120
cttaagtggg gatattatgta tttctcaagc aagtgattaa agcaaaacta ggcacgattg    180
aaatcaagat ctttttaggca anaaagtcac gatgagtttt agaattattt taggactctg    240
tggctttctc ttcatagaaa tagaaaaaaa aattgtataa aaccacaaaa ggtcctgaat    300
agccaaagca aactganca aaaagaacan agcaggaag caacacacta ccngaattca    360
aattatacta ccagggtgta gtaacaaaaa cagcattcta ttggcataaa atagacacca    420
agaccaatgg ancagaataa agaaccacac aaataaatcc atatatntac cgccanctga    480
ttatcaataa cnaacaccaa gaacatatnt taagggaacnt nctattcaat aantagtgt    540
ggnaaaaaact gggaaatcca tatgcagaaa naatgaaact agaccctat ccctcaccat    600
acgcaaannt caacttcgga atgggattac aaaacttaag acattccaac ccaagaaact    660
atnaaancta ctattaagaa aacagatcnc nccc                                694
```

<210> 666

<211> 705

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(705)

<223> n = A,T,C or G

<400> 666

```
tttaaaaaatt tagatacact angaaaatta ttttagtatac agaagaatat caggggggtgt    60
agtactcatc agagctaaat gagagcgctt taaaaatggt agtttgtctt ccgccatttc    120
tacagaaagc tgcaatttca ggttttcaac ctaatagggtg atatttaaga aaaaaaaaaa    180
gcaatcgcaa atagccccac tgcttttaca aatcattttt tctcttctag gtatagcctg    240
tcagggtggc taatgtaatt tttgacatct ctaggaattdt taatagaacc agaaatgggt    300
gccagagata tgctgcact aatcttaagt ggggattttat gtattttctca agcaagtgat    360
taaagcaaaa ctaggcacga ttgaaatcaa gatcttttag gcaagaaagt catgatgagt    420
tttanaatta ttttaggact ctgtggcttt ctcttcatag aaatagaaaa aaaaattgta    480
taaaaccaca aaaggctcctg aatagcccaa gcaacactga acaaaaagaa caaagcagga    540
agcaacacac taccagaatt caaattatac taccaagggtg tagtaaccaa aacagcattc    600
tattgggcnt aaaatagacc naagaccaat ggaacagaat aaagaacca aaataaatcc    660
atatttttac agccagctna ttatcaataa aaacnccaag aacnt                                705
```

<210> 667

<211> 817

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(817)

<223> n = A,T,C or G

```

<400> 667
nnangacttt tgtggtnnta tacaattntt ttttctattt ctatgaagag aaagccacag      60
agtcctaaaa taattctaaa actcatcatg actttcttgc ctaaaagatc ttgatttcaa      120
tcgtgcctag ttttgcttta atcacttgct tgagaaatac ataaatcccc acttaagatt      180
agtgcaggca tatctctggc acccatttct ggttctatta aaattcctag agatgtcaaa      240
aattacatta ggccacctga caggctatac ctagaagaga aaaaatgatt tgtaaaagca      300
gtggggctat ttgcgattgc tttttttttt tcttaaatat cacctattag gttgaaaacc      360
tgaaattgca gctttctgta gaaatggcgg aagacaaact aacattttta aagcgctctc      420
atthagctct gatgagtact acaccctga tattcttctg atactaaaat aattttccta      480
gtgtagtcta aactttttta aaaagacatg taatccgagg agtttgtaac tcaaaacgag      540
tgcacttagg aggtatcgca agccgtttct ggattaaatt ccagctagc ttgcttgctt      600
agcaggggag gnaaanaag acatctgcag cctagggag aaacaccttc gcattgttct      660
tacgtgttta cgttatttta tttcctanaa caaggcngaa ttgggactcg aatggttcag      720
ttgggggtggg ggatcccctg gtncataaaa ngtcanaaag anggtacagg cggaacncca      780
agggtcgtcc tgcatttana ctcggaatth ttggtgcc                               817

```

<210> 668

<211> 826

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(826)

<223> n = A,T,C or G

```

<400> 668
cgggggggnnt tacgtctctc tggacgcttt tattgtacca gggcgatccc agcccaactg      60
taccattcga gtcctactc ctgccttgct ctagggaaat aaaataacgt aaacacgtaa      120
gaacaatgag aaagcgtttt ctcccttagg ctgcagattg tcttcttcac cgccccgtct      180
tagctagcta gctagctggg aattttaatcc agaaacggct tgcgatacct cctagatgca      240
ctcgttttga gttacaaact ccgcggatta catgtctttt taaaaaagtt tagactacac      300
tagggaaaat tatttttagta tcagaagaat atcagggggg gtagtactca tcagagctna      360
atgagagcgc tttaaaaatg ttagtttgct ttccgccatt tctacagaaa gctgcaatth      420
caggttttca ncctaataag tgatatntaa gaaaaaaaaa acaatcgcan atagcccact      480
gctttttaca atcatttttc tcttctaggt atagcctgct aggtggccta atgtatthtt      540
gacatctcta ggaattttta tagaccagaa atgggtgcca gagatatgcc tgcactaatc      600
ttaagtgggg atttatgtat ttctcaanca agtgattaaa gcaaaactag gcacgaatga      660
aatcaagatc tttaggccag aaatcatgaa nanttttana attattttan gaatctgtgg      720
cttctcttct taaaatngaa aaaaaaattg tttaaaccca naagggtctga atacccaagc      780
nccctgaacn anagaacaan gccggagcac cccctcccaa atcccc                               826

```

<210> 669

<211> 547

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(547)

<223> n = A,T,C or G

```

<400> 669
cattgtgttg gggaaaaaat gatttgtata agcagtgggg ctatttgcca ttgctttttt    60
tttttcttaa atatcaccta ttaggttgaa aacctgaaat tgcagctttc tgtagaaatg    120
gcggaagaca aactaacatt tttaaagcgc tctcatttag ctctgatgag tactacaccc    180
ctnatattct tctgatacta aaataatttt cctagtgtag tctaaacttt tttaaaaaga    240
catgtaatcc gcggagttag taactcaaaa cgagtgcac tnggaagtat cgcagccgtt    300
nctggatnaa attcccagct tgctngcttg ctncgcccgg gggcggtnaa aaaaacatct    360
gcagcccngg ggnaaaaacc ttgcatttgt tcttacgtgt ttacgttatt ttatttcctt    420
nnagcaaggc nggganttg ggactcgaaa tggtagcgtt gggctgggga tcgcccttgt    480
tacataaaag ncgtccagaa gagggacggt tacaggcngg ganctccaaa ggtcagtcct    540
tgccatt                                         547

```

<210> 670

<211> 232

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1) ... (232)

<223> n = A,T,C or G

```

<400> 670
cgaactatct agactaccta ggaaaattat tttagtatca gaagaatata aggggtgtag    60
tactcatcag agctaaatga gagcgcttta aaaatgttag tttgtcttcc gccatttcta    120
cagaaagctg caatttcagg ttttcaacct aataggtgat atttaanaaa aaaaaaagc    180
aatcgcaaat agccccactg cttttacaaa tcattttttc cccaacacaa tg          232

```

<210> 671

<211> 214

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1) ... (214)

<223> n = A,T,C or G

```

<400> 671
ctccccctcc ntccttcgct actnencatt ttcnnaaatt tntttcgcnt atngggaaaa    60
acaccacat tnttcanctc gcacagaaca ngngggggtg tgtaaaatga agggcttccn    120
cnctttctct tattnaanaa cactnaaana gggangggct aaaaccgcg ngatntctac    180
nctatcgcg gcgcttttgg ngttggctag aaga                                         214

```

<210> 672

<211> 328

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature
 <222> (1)...(328)
 <223> n = A,T,C or G

<400> 672
 ngancagcgg ngtttaaacy ggctctaga ctcgaggaga cncctgttgg atggtggatc 60
 acanntcgt actactatac aggacagagt atcggganct cttggntggt ggngcctgcc 120
 aaccactgct nctgttaact gcgtatctga agggactcgg actggcttca gaagaactac 180
 cggctcgaat gnaccatgga tgattcncnc tagttgaaaa aaaactcagg cacatgtatt 240
 gccactgatg actagcgcca gactnctctc ggctctntaa cgagcccaca tgncngtgtg 300
 ncncctgtgc tgnctccaga agagggttc 328

<210> 673
 <211> 223
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(223)
 <223> n = A,T,C or G

<400> 673
 gggggcgaag ctggctagcg tttaaactta agcttggtac cgagctcgga tcccnagac 60
 attgtgcatg aaaatgcaaa ttgagtgtgg tctatantgc catctcacc tctgncngc 120
 tcaaaacaac ngctttctgc tgcaatgggt agggctcctn acncacggtc gcnnacggag 180
 gccncttat cctctcgggt nnggatccct ngaagcatnt tct 223

<210> 674
 <211> 256
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(256)
 <223> n = A,T,C or G

<400> 674
 gnggggtent ngatgagcgc gcgtaatacn atcactntcn ggcnngntgg gtaccgggcc 60
 cccctcnaa gcggccgcc ttttttntt ttttttcn acatgataan ntctttnttc 120
 taaacagacc acaccactan agttcctttn ctttngtacg gaattgagtt aaagtagagn 180
 atacaatgca gggcttcnnc tctatttcac attccaggnt ggttcngnat ggatcgcccc 240
 tgctctccg atgggt 256

<210> 675
 <211> 439
 <212> DNA
 <213> Homo sapien

<220>

<221> misc_feature
 <222> (1)...(439)
 <223> n = A,T,C or G

<400> 675
 nnactagtcc agtgtggtgg aattccattg tgttgggctt gtatggggtt ttttgtctag 60
 ttntttggga aatgttngtg ttactatntt ttggatatna tatatgatat gtatggccct 120
 tctatgggct cctcanacng aactcaacca ttttccacaa aaccnattcc tcctttccct 180
 tcatgactga gtggtgttgg tactatccng gaaactggga cattgtcctt cacatctntc 240
 ccttanactgc ctngtccnat tgatgtcttt gagctntgan atgtctttgt taactntctc 300
 ctntntctgt actgccggca naattaagca ccatntgtca caaaaagtat tgcgttacct 360
 tcacgnatct gttngttncc atncttgctg cttctccngn ggaaaatagg ctnttctggc 420
 aaccgaacng aanaaatac 439

<210> 676
 <211> 587
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(587)
 <223> n = A,T,C or G

<400> 676
 ngngggcctn attaagcgcg cgtaatacna ctcactntgg ggcgaattgg gtaccgggnc 60
 cccctcaagt tnatntgccn aacctctctt ttggaataac aaaagggttta acacatatgt 120
 cctcataggg acgcgctttc acacnttccg gacngcttca tanacntcat tntatattct 180
 cctcagnaca agttnaggen gaaggtgagg canacnttat aatttccatt tcacaaatnc 240
 ggaaagttag gctcaaaagg nttaaaaaat aacctgatac aantcataga gccggtntct 300
 ggaanaagca ggagcaaagt ccaggcatcc tgatccaagc tnggtccact gccttccact 360
 ctggagaggg ttcattctccg acaaaggaag ggacntgagt ggctgganaa tctcatggga 420
 taaagacctc agnatttcat gctcctggaa atcccatggg ttgaacaaca ggtntttggc 480
 ccgtggttct ntccctttgn ccatctttta accttggggg aaatgatggc ntctntnagc 540
 nttttttttn aaagagatng aaattgaatg attattngct cattggg 587

<210> 677
 <211> 444
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(444)
 <223> n = A,T,C or G

<400> 677
 gtggggcatn attaagcgcg cgtaatacga ctcactatag gggcgaantg ggtaccgggc 60
 cccctcgaa gcggccgccc tttttttttt tttttactgt ccaaactntc tatngatnta 120
 gttgaactgt ncaacgattt catgaaatcc tatacacana gccttcaggc ccagagagta 180
 aaacaaatth aaatttnttc accanattgn agcagncana agcatccnat natatccgac 240

tacaatgaat	natatgctna	nggtanctna	tttaccact	ntgggggtctt	tanggtctgt	300
cacaaactat	tttcgtaaac	atcnntttaa	anttnnggtga	atggacctaa	tnccagataa	360
ntctatttna	tntaccctag	catncctgtg	gctnactttt	cgggctgtgt	tggcntactt	420
ttaggagaaa	attggtataa	atnn				444

<210> 678

<211> 670

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(670)

<223> n = A,T,C or G

<400> 678

actagtccag	tgtggtggaa	ttccattgtg	ttgggagcag	tttaaaaaaa	aaaaagacna	60
aatatacnac	tcttgatnaa	acataaaggt	acagtgggtct	atgaggaana	gaaaaggtac	120
ctnaggatgc	aaaantacct	accacatggg	aaccgttngt	ccacactcat	tcnnananaa	180
accgagtcct	ctcanttnca	caogtgtacg	tttcagttgg	gaagtgcctg	ccattactcc	240
naagcctaga	accttcacgt	cctgaaggtt	ctggaaggtt	tttcagattg	cttaaganac	300
gcngcccttc	catattcntc	tcactaccc	nggggaacgg	aacaaatgga	gctgcgacng	360
ggaagcgtcc	cttcccntcc	gaacgctttc	tttcaaacct	gcctgccttc	cnggcgaatg	420
gaccggaagg	tttncnngct	tcctttcanc	ccnaattact	tcctgngttg	aaaattggcc	480
tgttggtttg	caaattcngg	aatttgttta	ctttcntcat	gtcctgtggt	gmnacnaaccg	540
gctcncctgt	tgccctccctt	tngaaaggtt	ttcatcaggc	cccgcctttt	ctcttntaan	600
ngtcctaate	cggncnggac	cactcgggga	aaattttttt	ttttcgaaaa	gccgccccnt	660
ccgtccggct						670

<210> 679

<211> 449

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(449)

<223> n = A,T,C or G

<400> 679

actagtccag	tgtggtggaa	ttccattgtg	ttgggagtag	gtctactaca	ncctacttcc	60
cctatcatan	aagancttan	caacnttcat	gatccccccc	tcntannoct	tttcctcanc	120
tgctccttag	tcctgtttgt	cctnttccca	acantcntaa	ganagatnac	taatnctact	180
atctctnacc	tcgggaanct	acaanaogtc	tggaactatt	cngaccccat	gcancncat	240
nctccatcgt	cctcccagcc	cctncccttc	ctttacntta	ctnaacgaag	gtcgacgatc	300
cctcccntac	ctcccnnncc	attgggnccc	aanggnactg	gacctcacga	ntacaccnac	360
tacgggnga	ctaagnctgn	aactccttac	atatntcccc	gttacccecn	gaacncagcg	420
aacngcnaca	ccttggaent	caagaanta				449

<210> 680

<211> 670

<223> n = A,T,C or G

<400> 682

tgatcattca	agcgntgngc	gnataacgat	tgctnagccc	aacctttcat	agggtcgttc	60
ctttgggaat	nggatgtcta	ttgaatggca	gggatagggg	cactcggcac	tcgcctctgg	120
tacagttttg	catatatatc	ctcatcgcca	gcgagcgtag	gggancgtta	agtttgggga	180
aatgccnccg	catgncctn	ccggagctta	aacccccaac	aatnccatt	ttnaaaaaag	240
ntttnttant	taaaaaaaaa	aac				263

<210> 683

<211> 255

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1) ... (255)

<223> n = A,T,C or G

<400> 683

cttgcccggc	atgcacagac	ntntttacgg	acacnctact	ccaagngagc	ctgnanctgt	60
ctacgggtcaa	nctctaaggt	tngncantgc	cacanatggc	atagtcccga	gggcggtnan	120
tctggantgc	tctctgcact	tgaacntaaa	gcgcntttca	aganaggnc	aatngcctgc	180
ctcttgacaa	cnaacaancc	cacaccnacc	tangaccctn	tangcaagga	ctggattctg	240
naaatgcaat	acaca					255

<210> 684

<211> 922

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1) ... (922)

<223> n = A,T,C or G

<400> 684

acctttcatt	tcattgtgctt	ctatttttcct	acatcttttta	catgactaag	ggattaatga	60
aatcacctct	tcataatcat	gaccataaatt	tcataccaaca	agtactcaag	tttggtgtta	120
gcacttttatt	aatgcttacg	aattctctct	ctctccctct	ttctcttttc	cttagtcctt	180
gcacaataag	gattttttgaa	tgtataatat	catcttaggt	aagctttcat	atggtttttg	240
catatgaagc	ttatgactgt	cataagccat	accaagcctg	tggagtatgg	catgattttc	300
attacataat	ccaatgaaaa	tagacttatt	ttaaattccct	aactttgtag	ttttaatttg	360
tatttcta	tcttgaaatt	aacagctagt	acttatccat	cacagcagtc	tcctactgac	420
atgaagcaag	ttgttgaaatg	cagtagancca	tgaatgaaag	catttaaatgt	tanacaaaaa	480
tgggtgatac	ccaagcattc	tgaattat	gcatacaagga	atgggacatg	tacattagtg	540
gcatactttc	taccaatatg	tgacttgaat	tgtttttttta	aaaaaaggan	aatgantttc	600
tcaatttgct	ttaaaaaaatt	ttnaaaaagt	tcaatggcat	gctgctttgt	ctggacttaa	660
tttattaaca	attnttaanc	cttccttaag	gacanaat	tggtgttcag	gatcnccttg	720
aagggtctta	tttttnatan	nattccaaac	ccaaaagggtg	gtttaaaatg	ggnggggttc	780
ccccncnaaa	atttggaccg	gcttttttat	atttaaaaaa	nttnccnttt	gngtttgaaa	840

```
nctnaatacc aattaagggg gaattttacc tnccagtggg aaaaaaaaaac nctngccntt 900
naaaaaattc ccnggagnca at 922
```

```
<210> 685
<211> 531
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(531)
<223> n = A,T,C or G
```

```
<400> 685
tgaggctctg taaaactgtt cctctgctag gcatacttca tattctctat attaaactca 60
tctttaattg gcatggaaga ttcattgttc caaatctcag atgaagatcc tatattggat 120
gcaattaagc ctggcagcgc cctcaaaaga cagtcttgtc actgctagcc acagccagga 180
cacagtaaca gttccttcta gtgacccnag accataanaa atananatct aaagaattct 240
gactccaaag gcattagccc attcctggta ttgccaatTA tgatagaaaa aattgccaaag 300
ctcctgggac atggaaatac actcagtaca tttgagaact ggagaactan tttccaaat 360
agtatgaaga catganggtg attgtagata tntgagtttg gagaanttga gggaaatcng 420
attacacatg tttactacaa gagatgttna taagtaaaga aggcctgata tacaatctaa 480
cagacnantg agataaatct taantcacia ctgaentccc ttttggggcg g 531
```

```
<210> 686
<211> 336
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(336)
<223> n = A,T,C or G
```

```
<400> 686
ggngncctna tgagcgcgcg taatacgatc atatagggcg aattgggtac cgggcccccc 60
tcaagaacac tacaagctat gtcctcttct canagagccc tgaantttta acatattgaa 120
agctctnatc ttgccaaana actccactta acttcaaaac acaccctcca cacacatcat 180
gatcaactna gatcttactg aaccagaatc ctnaatggca tacttcagga acaggggtcc 240
anagaagcag ttctcaaant gcagctnaaa aagaaactga aaaccaatt catgcaanac 300
ctagggctta tttgagagca ttttcagtg cagatt 336
```

```
<210> 687
<211> 271
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(271)
<223> n = A,T,C or G
```

```

<400> 687
aatctgcact ggaaaatgct ctaaaataag ccctaggtct tgcatagaatt gggttttcag      60
tttcttttta agctgcactt tgagaactgc ttctctggac ccctgttcct gaagtatgcc      120
atntagatt ctggttcagt aagatctcag ttaatcatga tgtgtgtgga ggggtgtgtt      180
tgaagttnag tggagttctt tggcaagatc agagctttca atatgttnaa acttcagggc      240
tctctgagaa gaggacatag cttgtagtgt t                                271

```

<210> 688

<211> 740

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(740)

<223> n = A,T,C or G

```

<400> 688
tgatgaagcg cgcgtnntac nactcactat nggggcgaan tatgggtacc gggnccccct      60
cgaagcggcc gccctttttt tntttttttg tgagagttta aataaaatat ttgagtttaa      120
tttaaagttt gagtttaatt aaaatatatg gcatatccca agttgggctt tgcanaaaga      180
acacttctca ggaactgtta gttggtgtac caggaactca gaagggtcct gttattaaat      240
atatttgtaa aatgcatgga ttctctgaan atcnctctgc atgtgagcaa cacttacatc      300
ncaaaccaaa attggcattg catacatnaa ccaatatttc ccaaacattt ctggttatgg      360
cccaccccct ttgtgtanta cttattgctg ttttttgtaa ccctggggaa attacttaaa      420
atattcagct ggaaattaca ggcgttactt ttaaggganc aagaattaca gtgactccca      480
aaattgcaag tgttgattac tatttaagaa cccaagaatt tgaaagaaat tttgaaaagt      540
gaaaacngga aatnttaaat gacttctcaa attttgaaaa ctcnngnaaa catctccact      600
ttggtncctt tccttttaaaa attggctaaa aattntttnt tatncccacc ccattggaan      660
tncccccccc ctggaacaat tggattcccc tatttcctaa aaaacggccn ccccccccg      720
ggngaacncc nacnttttgn                                740

```

<210> 689

<211> 635

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(635)

<223> n = A,T,C or G

```

<400> 689
actagtccag tgtgggtgga ttccattgtg ttgggattac atatactttt agcaattttt      60
aaagaagtgt acaaagttga gatgtttcct gagctctcat atatctgana atgtcatttt      120
acatctccgt cttcacctct caaaacttct ttcaattctt tggctcttaa tagtaatcaa      180
cacttgcact ctggagtcac tgtaattctt gctcctttac agctacnctt gttatttcca      240
gctgaatatt tttagttatt tcccaggggt ccaaaaaaca gcaataagta ctacacaaag      300
gggggtgggcc ataaccagaa atgtttggga aatactggct catgtatgca atgccaaatc      360
tggtttgcna ttgtantgtt gctcacatgc agagtgaatc ttcaaanaat ccatgcattt      420

```

tccaaatata	tttaataaca	gggaaccttc	tganttctctg	gntacaccaa	ctaacagttc	480
ctgaaaaaatg	ttcttttctgc	aaaacccaac	ttggggatat	gccatatatt	ttaattaaac	540
tcaaaacttta	aattaaactn	caattatttt	atttttaaact	cctcaaaaaa	aaaaaaaaaa	600
aggggggggcc	cttccaangg	gggggnccggt	tcccc			635

<210> 690

<211> 3923

<212> DNA

<213> Homo sapien

<400> 690

acagaagaaa	tagcaagtgc	cgagaagctg	gcatcagaaa	aacagagggg	agattttgtgt	60
ggctgcagcc	gagggagacc	aggaagatct	gcatggtggg	aaggacctga	tgatacagag	120
gaattacaac	acataacttt	agtgtttcaa	tgaacaccaa	gataaataag	tgaagagcta	180
gtccgctgtg	agtctcctca	gtgacacagg	gctggatcac	catcgacggc	actttctgag	240
tactcagtgc	agcaaagaaa	gactacagac	atctcaatgg	caggggtgag	aaataagaaa	300
ggctgctgac	tttaccatct	gaggccacac	atctgctgaa	atggagataa	ttaacatcac	360
tagaaacagc	aagatgacaa	tataatgtct	aagtagtgac	atgtttttgc	acatttccag	420
cccctttaaa	tatccacaca	cacaggaagc	acaaaaggaa	gcacagagat	ccctgggaga	480
aatgcccgcc	cgccatcttg	ggtcacgat	gagcctcgcc	ctgtgcctgg	tcccgcttgt	540
gaggaagga	cattagaaaa	tgaattgatg	tgttccttaa	aggatgggca	ggaaaacaga	600
tcctgttgtg	gatatttatt	tgaacgggat	tacagatttg	aatgaagtc	acaaagttag	660
cattaccaat	gagaggaaaa	cagacgagaa	aatcttgatg	gcttcacaag	acatgcaaca	720
aacaaaatgg	aatactgtga	tgacatgagg	cagccaagct	ggggaggaga	taaccacggg	780
gcagagggtc	aggattctgg	ccctgctgcc	taaactgtgc	gttcataacc	aaatcatttc	840
atattttctaa	ccctcaaaac	aaagctgttg	taatatctga	tctctacggg	tccttctggg	900
cccaacattc	tccatatatc	cagccacact	cattttttaat	atthagttcc	cagatctgta	960
ctgtgacctt	tctacactgt	agaataacat	tactcatttt	gttcaaagac	ccttcgtgtt	1020
gctgcctaatt	atgtagctga	ctgtttttcc	taaggagtgt	tctggcccag	gggatctgtg	1080
aacaggcttg	gaagcatctc	aagatctttc	caggggtata	cttactagca	cacagcatga	1140
tcattacgga	gtgaattatc	taatcaacat	catcctcagt	gtccttgccc	atactgaaat	1200
tcattttccca	cttttgtgcc	cattctcaag	acctcaaaat	gtcattccat	taatatcaca	1260
ggattaactt	tttttttttaa	cctggaagaa	ttcaatgtta	catgcagcta	tgggaattta	1320
attacatatt	ttgtttttcca	gtgcaaagat	gactaagtcc	tttatccctc	ccctttgttt	1380
gatttttttt	ccagtataaa	gttaaaatgc	ttagccttgt	actgaggctg	tatacagcac	1440
agcctctccc	catccctcca	gccttatctg	tcatcaccat	caaccctcc	cataccacct	1500
aaacaaaatc	taacttgtaa	ttccttgtaac	atgtcaggac	atacattatt	ccttctgcct	1560
gagaagctct	tccttgtctc	ttaaactctag	aatgatgtaa	agttttgaat	aagttgacta	1620
tcttacttca	tgcaaagaag	ggacacatat	gagattcatc	atcacatgag	acagcaaata	1680
ctaaaagtgt	aatttgatta	taagagttaa	gataaatata	tgaaatgcaa	gagccacaga	1740
gggaatgttt	atggggcacg	tttgtaagcc	tgggatgtga	agcaaaggca	gggaacctca	1800
tagtatctta	tataatatac	ttcatttctc	tatctctatc	acaatatcca	acaagctttt	1860
cacagaattc	atgcagtgca	aatccccaaa	ggtaaccttt	atccatttca	tgggtgagtgc	1920
gctttagaat	tttggcaaat	catactgggc	acttatctca	actttgagat	gtgtttgtcc	1980
ttgtagttaa	ttgaaagaaa	tagggcactc	ttgtgagcca	ctttaggggt	cactcctggc	2040
aataaagaat	ttacaaagag	ctactcagga	ccagttgtta	agagctctgt	gtgtgtgtgt	2100
gtgtgtgtgt	gagtgtacat	gccaaagtgt	gcctctctct	cttgacccat	tatttcagac	2160
ttaaaacaag	catgttttca	aatggcacta	tgagctgcca	atgatgtatc	accaccatat	2220
ctcattattc	tccagtaaat	gtgataataa	tgtcatctgt	taacataaaa	aaagtttgac	2280
ttcacaaaag	cagctggaaa	tggacaacca	caatatgcat	aaatctaact	cctaccatca	2340
gctacacact	gcttgacata	tattgttaga	agcacctcgc	atthgtgggt	tctcttaagc	2400

```

aaaataacttg cattaggtct cagctggggc tgtgcatcag gcggtttgag aaatattcaa 2460
ttctcagcag aagccagaat ttgaattccc tcatctttta ggaatcattt accaggtttg 2520
gagaggattc agacagctca ggtgctttca ctaatgtctc tgaacttctg tccctctttg 2580
tgttcatgga tagtccaata aataatgtta tctttgaact gatgctcata ggagagaata 2640
taagaactct gagtgatata aacattaggg attcaaagaa atattagatt taagctcaca 2700
ctgggtcaaaa ggaaccaaga tacaaagaac tctgagctgt catcgteccc atctctgtga 2760
gccacaacca acagcaggac ccaacgcatg tctgagatcc ttaaatcaag gaaaccagtg 2820
tcatgagttg aattctccta ttatggatgc tagcttctgg ccctctctgg ctctctctct 2880
gacacatatt agcttctagc ctttgcttcc acgactttta tcttttctcc aacacatcgc 2940
ttaccaatcc tctctctgct ctggtgcttt ggacttcccc acaagaattt caacgactct 3000
caagtctttt cttccatccc caccactaac ctgaatgcct agacccttat ttttattaat 3060
ttccaataga tgctgcctat gggctatatt gcttttagatg aacattagat atttaaagct 3120
caagaggttc aaaatccaac tcattatctt ctctttcttt cactctcttg ctctctctcc 3180
tatattactg attgcactga acagcatggg cccaatgta gccatgcaaa tgagaaaccc 3240
agtggctcct tgtgtacat gcagcaaga ctgctgaagc cagaaggatg actgattacg 3300
cctcatgggt ggaggggacc actcctgggc cttcgtgatt gtcaggagca agacctgaga 3360
tgctccctgc cttcagtgtc ctctgcatct cccctttcta atgaagatcc atagaatttg 3420
ctacatttga gaattccaat taggaactca catgttttat ctgccctatc aattttttaa 3480
acttgctgaa aattaagttt tttcaaaatc tgtccttgta aattactttt tcttacagtg 3540
tcttggcata ctatatcaac tttgattctt tgttacaact tttcttactc ttttatcacc 3600
aaagtggctt ttattctctt tattattatt attttctttt actactatat tacgttggtta 3660
ttattttggt ctctatagta tcaatttatt tgatttagtt tcaatttatt tttattgctg 3720
acttttaaaa taagtgattc ggggggtggg agaacagggg agggagagca ttaggacaaa 3780
tacctaattg atgtgggact taaaacctag atgatgggtt gataggtgca gcaaacact 3840
atggcacacg tatacctgtg taacaaacct acacattctg cacatgtatc ccagaacgta 3900
aagtaaaatt taaaaaaaag tga 3923

```

<210> 691

<211> 882

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1) ... (882)

<223> n = A,T,C or G

<400> 691

```

ttactcacta tagggctcga gcggccgctg aattctgctg cagtgagctg tgattatgtc 60
cctgcactcc agcctggatg acagaacacg atcattttctc taaagacaaa caaaaaacat 120
aaaataaaac tagtataagg atagaagccc agggttgatt taagtctgcg gaaatcataa 180
accataggtc agacttctca ttgatgaggt acttggtgggt tagaatacaa ttaggtatat 240
ttggtctaga aaccaggatg gaattagaga ataaaagact gagcaatagc atgttatagt 300
attagaaata ctatagaaat aggaaaagcc ctgattatga ctttggagtt ctgatccaac 360
atctgggatt atttagatat tttaaaggaa aacgatgact ttttagctctc aggatgttag 420
tttctcctca cataaaatga agagcctcga aaagatttctg tttaccagat tatttctgaa 480
gtcaattcca gttctaaaat tccatcactg ngcactaagg caaattgaat tgaataaagt 540
attgggnatg cataaaatag tctattttta aaaangaata gtaattatcc attggnaaca 600
gacgcantca tccagncatc tcttacctg ncccatgnen tatgtagana tgtanctcta 660
atcccttaac aaaccgattt tgcaaaggag cttanccttg gggtaacttg tcanggcaac 720
tgggtctact tnaagactca tcttcaacta ctgggcacca aatncctacc attgcatcaa 780

```


actgggggttc ccatncaagg caaacccctgn gaaatcttta atcccgaat tggcgcccaa	840
ttttgngggg tttccnaaaa gaatcntccc ccccgagggg cc	882

<210> 692

<211> 235

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(235)

<223> n = A,T,C or G

<400> 692

ccgcactngt aangnccgcc agngngctgn aantccgctn agcncggatc cactagtcca	60
ttgatggtaa aagggtagct tactggnatg tccgctgct ccanganata atacncagga	120
cttctcanag cacttaatat gttaataata aactncngna aaaaagatnt tcnatgaanc	180
nttcctctta ggaggtcagg ngagaatagt gttaatgnca ttaagganag aacga	235

<210> 693

<211> 383

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(383)

<223> n = A,T,C or G

<400> 693

nttatgtaag aaatgtcata tatcttttat tttctttaaa tcaaaataaa tatgactttg	60
agcatcccat cccatgcccc atcctatcag aatggtagga acatcaacac aaataattag	120
taatgcaccg catctacatt cccatgctct ctttacttct tcagcattgc ctaaaggcat	180
aatacacctt taattaatta attcagcctc ctaatgcaca ttaacaaagc ccctgctaga	240
ctctgtccat aatggnaaac ctgnatgatc cttgatatta acantttaag gaatgctcat	300
ggattggtn cagacttaaa aaattgaggg ggctgaanaa aatctaangg anaaatcatg	360
gaagcatttg cacatattac ata	383

<210> 694

<211> 204

<212> DNA

<213> Homo sapien

<400> 694

tctcttggtt ggtcagcctg aagggtggta atgactcacc aacgctacta atccttcttc	60
actgtccctt atttttttcc ctcccaggct cataactcga ggttaaactc tcttttatac	120
aagaacctg tctgatgaag catcatttca gaattttaag tcaacttaca aatgtggtat	180
tattcacatc tgagtacaaa tttta	204

<210> 695

<211> 670

ctncagctct	aatcgactnc	tatnaggnat	gatggcncgt	gcngcgcgta	cgtantgctt	60
ggatcctcnn	anagcggacg	cctactacta	ctaaattcgc	ggncgcggtg	actttttttg	120
tttttttccct	tnacagagnt	ntttttgtgc	ccttggttct	tatgctcana	ctcngcaaaa	180
aanatcaaaa	gntacnnatg	aaaaacntat	nccatctnca	naaaggaggt	gnagntatta	240
ctttct						246

<210> 698

<211> 3674

<212> DNA

<213> Homo sapien

<400> 698

agaaagtttc	cttttttttt	tttaatggtg	aaaagatata	cacatattta	gaattagcca	60
gctgggctca	gtttagatta	ttccaatttt	gttggcaaca	tccagagcat	cgtaatcagg	120
agccagtga	acatattcct	tcttctctcc	atcaggccaa	atcacggtgt	tgacctggc	180
cacatcaatg	tcttagaact	tcttcacagc	ctgtttgatc	tggtgcttgt	tggtcttaac	240
atccacaatg	aacacaagtg	tgttgttgtc	ttctatcttc	ttcgtggtga	ctcagtggtc	300
agcggaaact	tgatgatagc	gtagtggtea	agcttgtatc	tcctgggagc	gctcttccaa	360
agatatttgg	gctgcctcgg	gagttgcagc	gtcttgggcc	gccggaaggt	gggtgacgta	420
cggatcttct	ttttttgtgt	ggctgtggac	acctttcaac	actgtcttct	tggtctttaa	480
atccttcgct	ttggtttcgg	ctataggagg	ggcaggagct	tccttcttca	ctttcggcgc	540
catcttgtga	aaagggaag	tttcttttct	aataccattt	tcacttctcc	cgaattttgt	600
ggatcgtttc	ttggtatcta	ccccagattt	caggagtgtt	ggctggatct	tagggattgt	660
gaagtcttca	tttccctgtg	gtgagatctg	aggcatgatt	ttaaacagtg	tgagggaagg	720
agatctccag	gcactttaat	agaatggaga	agcaggatgg	gatttgagag	gaaatctgat	780
tttgaaaaaa	ggagaactag	agttgagttc	gtaattaact	agcaccttaa	aggtcattca	840
gcatgcccat	ctgcacagtg	ggtgtaatca	ccctacagaa	caaaaacaaa	aaggcaatgg	900
agaggaagct	gtaaagcact	gtacatgttt	aactcattgt	tatgtaagct	agccgaaggc	960
ttcacagact	tgaattcatc	tcccagtttc	tcttctgtga	ctggaaactc	tgcttaggt	1020
tgcttaaaac	ttgagaaaaca	gaatattgct	tcccctgcct	gccttcttga	gtacacttgc	1080
ctacacaaag	atgcacatcc	ttgtttgtgt	gtgtgtgtcc	atttgctgtg	acattcttgt	1140
gaaagtcaaa	gtttcccagc	tgttgacata	cacaagtttg	tttggtgcaa	cctgtcagat	1200
gcatccctta	gacaggccct	ttgatactct	gggaaagaca	ttggacttac	agtcggaacg	1260
aaaagaaaag	aatgtgatat	gtatagcgtg	cagtgagttg	gagttttacc	tgtattgttt	1320
taatttcaac	aagcttgagg	actagccaca	aatgtaccca	gtttacaaat	gaggaaacag	1380
gtgcaaaaag	gttgttacct	gtcaaaggte	gtatgtggca	gagccaagat	ttgagcccag	1440
ttatgtctga	tgaacttagc	ctatgctctt	taaacttctg	aatgctgacc	attgaggata	1500
tctaaactta	gatcaattgc	attttccctc	caagactatt	tacttatcaa	tacaataata	1560
ccacctttac	caatctattg	ttttgatacg	agactcaaat	atgccagata	tatgtaaaag	1620
caacctacaa	gctctctaata	catgctcacc	taaaagattc	cgggatcta	ataggctcaa	1680
agaaacttct	tctagaaaata	taaaagagaa	aattggatta	tgcaaaaatt	cattattaat	1740
ttttttcatc	catcctttta	ttcagcaaac	atztatctgt	tgttgacttt	atgcagtatg	1800
gccttttaag	gattggggga	caggtgaaga	acgggggtgc	agaatgcac	ctcctactaa	1860
tgaggtcagt	acacatttgc	attttaaaat	gccctgtcca	gctgggcatg	gtggatcatg	1920
cctgtaatct	caacattgga	aggccaaggc	aggaggattg	cttcagccca	ggagttcaag	1980
accagcctgg	gcaacataga	aagaccccat	ctctcaatca	atcaatcaat	gccctgtctt	2040
tgaaaataaa	actctttaag	aaaggtttaa	tgggcagggt	gtggtagctc	atgcctataa	2100
tacagcactt	tgggaggctg	aggcaggagg	atcactttag	cccagaagtt	caagaccagc	2160
ctgggcaaca	agtgcacct	catctcaatt	ttttaataaa	atgaatacat	acataaggaa	2220
agataaaaag	aaaagtttaa	tgaaagaata	cagtataaaa	caaactctct	ggacctaaaa	2280
gtatttttgt	tcaagccaaa	tattgtgaat	cacctctctg	tgttgaggat	acagaatatc	2340

taagcccagg	aaactgagca	gaaagtccat	gtactaacta	atcaacccga	ggcaaggcaa	2400
aaatgagact	aactaatcaa	tccgaggcaa	ggggcaaatt	agacggaacc	tgactctggg	2460
ctattaagcg	acaactttcc	ctctgttgta	ttttctttt	attcaatgta	aaaggataaa	2520
aactctctaa	aactaaaaac	aatgtttgtc	aggagttaca	aaccatgacc	aactaattat	2580
ggggaatcat	aaaatatgac	tgtatgagat	cttgatgggt	tacaaagtgt	acccactggt	2640
aatcacttta	aacattaatg	aacttaaaaa	tgaatttacg	gagattggaa	tgtttctttc	2700
ctgttgatt	agttggctca	ggctgccata	acaaaatacc	acagactggg	aggcttaagt	2760
aacagaaatt	catttctcac	agttctgggg	gctggaagtc	cacgatcaag	gtgcaggaaa	2820
ggcaggcttc	attctgaggc	ccctctcttg	gctcacatgt	ggccaccctc	ccactgctg	2880
ctcacatgac	ctctttgtgc	tcttggaag	aggggtgtggg	ggacagaggg	aaagagaagg	2940
agagggaaact	ctctggtgtc	tctgttttca	aggacctaa	cctggggccac	tttggccag	3000
gcactgtggg	gtgggggggt	gtggctgctc	tgctctgagt	ggccaagata	aagcaacaga	3060
aaaatgtcca	aagctgtgca	gcaaagacaa	gccaccgaac	agggatctgc	tcatcagtgt	3120
ggggacctcc	aagtcggcca	ccctggaggc	aagcccccam	agagcccatg	caaggtggca	3180
gcagcagaag	aagggaattg	tccctgtcct	tggcacattc	ctcaccgacc	tggtgatgct	3240
ggacactgcg	atgaatggta	atgtggatga	gaatatgatg	gactcccaga	aaaggagacc	3300
cagctgctca	ggtggctgca	aatcattaca	gccttcatcc	tggggaggaa	ctgggggcct	3360
ggttctgggt	cagagagcag	cccagtgagg	gtgagagcta	cagcctgtcc	tgccagctgg	3420
atccccagtc	ccggtcaacc	agtaatcaag	gctgagcaga	tcaggcttcc	cggagctggt	3480
cttgggaagc	cagccctggg	gtgagttggc	tctgtctgtg	gtactgagac	aatattgtca	3540
taaattcaat	gcgccttgt	atcccttttt	cttttttata	tgtctacatc	tataatcact	3600
atgcatacta	gtctttgtta	gtgtttctat	tcmacttaat	agagatatgt	tataacttaa	3660
aaaaaaaaaa	aaaa					3674

<210> 699

<211> 2051

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1) ... (2051)

<223> n = A,T,C or G

<400> 699

ggaccagggg	ctgaagtga	ccccagcac	agcacagctg	ctctataaaa	acgtggccag	60
actttttttt	ttgaagcaag	tccctgttct	tgttcgtcct	gactagtccc	atcagggcc	120
tggatcccaa	gactcagcat	ccaaggtccc	ctccaggaat	cctggcagct	cagcatactt	180
tatcctgttt	catctgagag	caaaaatgta	aaattggatg	cacagaaaag	tgactcaaag	240
tgcttaatga	ctagaagaaa	tctaggagca	gcaagaagag	caggacaaac	aggccaggcg	300
gtgtcaggag	cccagggtctc	cagctggang	gaacgtcaac	cctgcagtgg	gagcaggggc	360
cctttgcaca	tcttaggcac	agatggtaat	gtagacacca	caggtaagct	gggcttggt	420
cctacccttc	cccgattca	gaaagaaacc	aaacaaggag	ctttgtgtgg	aatgaaacct	480
cctttcctcc	cagaagcact	gctgactggt	tggtggttgc	catttgtggc	agtgagccct	540
tgtttgttct	gaggttgggc	tggtttctcc	tcttggccct	gccctacaga	tcataaagga	600
gaacagcaag	acgtccccag	caaacatcca	cagatggcct	tggaaataag	tcaccttct	660
caccctgcag	gaatgccagt	gaacatattg	ctgacatctt	ggagctcagt	acctcatagt	720
gtaacggcgt	cagtagatct	gcctgtgctg	ggacttccct	tactacccat	tctgagggg	780
cgatgcttct	gcagggcctg	tgacttggtg	cacaacttca	gacaccatca	tcttgagca	840
gcaccgcacc	ctcactagcc	aggggtgttg	tgacttccct	aaggccaagg	ccacattcaa	900
ggcttcggac	ttcattgatg	cgcttgtgct	gagcaagggt	gcttctccgg	gatcttaatt	960

ggactgtacc	atatcagtgg	agagctgcag	caaggtggcc	cctagagcca	cgcaccagcc	720
tgcacattgc	ctctccatac	ggcagccctt	tatttgga	cttcctaaat	cactttgctg	780
tgtgtgttta	cacgggtgtg	ttttgcttta	cttgccctga	gagcacacgg	gagtgcagca	840
cacaccccaa	cccacatcaa	ctgccattaa	agaaaagaaa	tttcagccca	gaatttcctg	900
tccagcaaaa	ttaagcatca	taagtgaagg	agaaataaga	tccttttcag	acaagcaagt	960
gctgagggaa	tttggtatca	ccagatctac	cttacgagag	ctcctgaagg	aagcactaaa	1020
tatggaaaga	aaagatcatc	acctgctact	acaaaaacac	actgaagtac	acagtccaat	1080
gatgctaaaa	agcaagcaca	tatgtaagtc	tgcaaaataa	ccagctgaca	gcatgacgac	1140
aggataaaat	ccacacatac	cattactaac	cttaaagtga	aatgggctaa	atgctcccat	1200
tgaaagacac	ggggcaagct	gggtaaagaa	ccaagaccca	ctggagtatg	ccgtcttcaa	1260
gcaacccatc	tcacgtgcag	tgccatacat	aggctcaaaa	taaaggaatg	gagaaaaata	1320
tttcaagcaa	atggaaaaca	gaaaaaaggt	gttgactctc	cagtttctga	caaacagac	1380
tctaccaata	aagataaaaa	aagagaagga	cattacaaag	gtggctctga	cctttgataa	1440
atctcattat	tgcttgatac	caacctgggc	tatttgtatt	gccccaacga	ataggataat	1500
ttgctgaggt	tgtggagctt	ctccccttca	cagagtccct	gatctccgaa	aatttggttg	1560
agatgtaagg	tgtattttgc	tgtacaactc	cttttttgaa	gttttactca	tttccaacaa	1620
ggaaggcaag	ttttcctgct	tccattgaca	aaggagagca	ggcacctcct	ttcctgagtt	1680
tcagcttgct	tctgacaggg	aaggagcttt	gagatttgaa	tactggcctg	ctgggttttg	1740
gacgtgcatt	gggcctgtgg	tcccatttgt	gttatttttc	tgggaaattt	cttccctttg	1800
gagtggagaa	gcttacccaa	tgctgtgacc	atcatcgtac	cttaaaagaa	ctccatttta	1860
agttcaggga	ctccttggca	gaagagaccg	tagccttgta	tcagatcata	aaggagaaga	1920
gcaagaggtc	cccggaac	atccacagat	ggccttggaa	ataagtcacc	ttgctcaccc	1980
tgagggaatg	ccagtgaact	tattgctgac	atcttggagc	tcagtaccct	catagtgtaa	2040
cggcgtcagc	agatctgcct	gtgctgggac	ttcctgtact	acctattcct	gaggggcat	2100
gcttctgcag	ggcctgtgac	ttggtgcaca	acttcagaca	ccatcatctt	gcagcagcac	2160
cgcaccctca	ctagccaggg	tgttgatgac	ttcctcaagg	ccaaggccac	attcaaggct	2220
tcggacttca	ttgatgcgct	tgtgctgagc	aaggtggctt	ctccgggatc	ttaattcagg	2280
aggtagaatg	gagcttgaga	tcaagtgtct	gatcaagcct	cagtgtatgg	gcgctgttca	2340
tcntctggty	ctgaagcagc	caagagaccc	aagtctgcct	ggctgcntct	taggatatga	2400
cagcagagcc	agtggcctct	actagatcct	gtacaacctc	acaaaacacc	cagacatcgg	2460
gagtgcctgc	agcctgtgat	gcaagagtc	taatcctgaa	gacattgaat	gacctgtcat	2520
tctgctgttt	ttacccaaaa	ggatcatgag	gatcagagag	gaaaagtcac	ttgcccaaag	2580
tcacacagct	gaacagtgg	ggagttcaac	tttgaccgtg	ggctgtctga	ccccagggtg	2640
tatgcttgct	tctctcccaa	gagacaactt	tcttatcagg	ctcaaataga	tgaaaggagg	2700
atgttaaagg	taggatctct	gaagcctgtg	ccagtggaa	cgcagctcat	ggctggcacc	2760
tgtgttctca	ttcttacctc	attaagagta	aagtttattg	agtttattga	atttaagtat	2820
ctttagttag	atcatatatt	attagtaaga	actgggacca	aacagatttt	ctgactctaa	2880
aagagagatt	ttcacagaaa	cagatatata	cctgtaagta	tacagacacg	catacacaca	2940
tttctttact	gctcataaaa	attagtcctt	attagaatgt	gggatgtata	aatgtaagag	3000
aattttcatg	ttaaaattga	cagatacatt	tttaaatgtg	cctaaaataa	atttaattat	3060
ttttntttta	gaattttcca	ttattaatgt	tatttttatg	agaaactata	taacttttatt	3120
gataatacat	acaataaccc	tttgtttttc	aaattgaaaa	tacagtgtat	tttgcaataa	3180
actaagtcct	aattttgtat	taaaatttta	aattttcaaa	aaaaaaaa		3228

<210> 702

<211> 4894

<212> DNA

<213> Homo sapiens

<400> 702

ctgcacgcgc tggctccggg tgacagccgc ggcctcggc caggatctga gtgatgagac 60
 gtgtccccac tgaggtgccc cacagcagca ggtgttgagc atgggctgag aagctggacc 120
 ggcaccaaag ggctggcaga aatgggcgcc tggctgattc ctaggcagtt ggcggcagca 180
 aggaggagag gccgcagctt ctggagcaga gccgagacga agcagttctg gagtgcctga 240
 acggccccct gagccctacc cgcttgcccc actatgggtc agaggctgtg ggtgagccgc 300
 ctgctgcggc accggaagc ccagctcttg ctggccaacc tgctaaccct tggcctggag 360
 gtgtgttttg ccgcaggcat cacctatgtg ccgcctctgc tgcctggaagt gggggtagag 420
 gagaagtcca tgaccatggt gctgggtgag tcaactacat ctccttccct cctgttccag 480
 atacatgcc cctggcatgt gggacaggag tacctctgcc ctgggagctg cttggaggga 540
 gaggtggtct gctgggaagg cattgtctgg caggagggtg accctgggct gagggggcac 600
 accaagagaa agaagagaat accaaggaca taccacagtc accctctggat ccctggctct 660
 gcacagagcc tggctcatag gagacactgg agaaatgctc ctaacccttg gctagccctt 720
 ttataattta tagcgattat ctcatttaat gcttacaacc accatttgag gtgatccatt 780
 ttacagagaa ggaagcagag gcttttaaga ggttaggtaa gtcttagcca aagccaaata 840
 gcagctgaac agtagagctg ggactccatc aaggctctcc agccggagct tgctcctacc 900
 cctaggacaa ggggtggact cctgactctg cagataaatt ctacaaaagc cacagaaggc 960
 aagtagtaac cattgtgtga caaccctca ccccaggaa gagggggccc tgtgaggatt 1020
 gcaggctctg ggtcacact gctgtgtgaa acgctgcctc ttaccctccc taggtctgcg 1080
 cctttgaata agtatcactt cttagttgct ccatgcctca gtttgtccat ctgaaaatgg 1140
 gggcatctgt aatgcctgtg ttatgaggag taaattacag catcctctgt aagacgtagc 1200
 acagtgtcga gtacggaatg ttatttccat ccttctcacg gagcttggtt ccccttcccc 1260
 ttgcccttta cttgtcccag ccattgactc atactacttc ccttcttgca ggcattgggtc 1320
 cagtgtctgg cctggtctgt gtcccgctcc taggtcagc cagtgaccac tggcgtggac 1380
 gctatggccg ccgcggccc ttcatctggg cactgtcctt gggcatcctg ctgagcctct 1440
 ttctcatccc aagggccggc tggctagcag ggctgtctgt cccggatccc agggccctgg 1500
 agctggcact gctcatcctg ggcgtggggc tgcctggact ctgtggccag gtgtgcttca 1560
 ctccactgga ggccctgctc totgacctct tccgggaccc ggaccactgt cgcaggcct 1620
 actctgtcta tgcctcatg atcagtcttg ggggtgcctt gggctacctc ctgcctgcca 1680
 ttgactggga caccagtgcc ctggccccct acctgggcac ccaggaggag tgcctctttg 1740
 gcctgtcac cctcatcttc ctacactgcg tagcagccac actgctggtg gctgaggagg 1800
 cagcgtctgg cccaccagag ccagcagaag ggctgtcggc cccctccttg tcgccccact 1860
 gctgtccatg ccgggcccgc ttggctttcc ggaacctggg cgccctgctt ccccggtgc 1920
 accagctgtg ctgccgcatg ccccgccccc tgcgcgggct cttcgtggct gagctgtgca 1980
 gctggatggc actcatgacc ttacagctgt ttacacgga tttcgtggg gaggggctgt 2040
 accagggcgt gccagagct gagccgggca ccgaggccc gagacactat gatgaaggta 2100
 aggccttggc agccagcaga ggctggtgtg ggagccggc accagagacg acactcgggg 2160
 ctgtgtctgg gctggtgcct ctccatcctg gcccgactt ctctgtcagg aaagtgggga 2220
 tggaccccat ctgcatacac ggcttctcat ggggtgtgaa catctctgct tgcggtttca 2280
 ggaaggcctc tggctgctct aggagtctga tcagagtcgt tgccccagtt tgacagaagg 2340
 aaaggcggag cttattcaaa gtctagaggg agtggaggag ttaaggctgg atttcagatc 2400
 tgctgtgttc cagccgcagt gtgccctctg ctcccccaac gactttccaa ataactctac 2460
 cagcgccttc cagctcaggc gtccatagaag cgtcttgaag cctatggcca gctgtctttg 2520
 tgttccctct caccgcctg tccctcacagc tgagactccc aggaaacctt cagactacct 2580
 tccctgctt cagcaaggg gcgttgccca cattctctga gggtcagtgg aagaacctag 2640
 actcccattg ctagaggtag aaaggggaag ggtgctgggg agcagggtct gtccacagca 2700
 ggtctcgtgc agcaggtagc tgtggttccg ccttctcatc tccctgagac tgctccgacc 2760
 cttccctccc aggtctgtgc tgatggcccc tctccctctg caggcgttcg gatgggcagc 2820
 ctggggctgt tccctcagtg cccatctcc ctggtcttct ctctggtcat ggaccggctg 2880
 gtgcagcgat tcggcactcg agcagctctat ttggccagtg tggcagcttt ccctgtggct 2940
 gccggtgcc catgcctgtc ccacagtgtg gccgtggtga cagcttcagc cgccctcacc 3000
 gggttcacct tctcagccct gcagatcctg ccctacacac tggcctccct ctaccaccgg 3060

gagaagcagg tgttcctgcc caaataccga ggggacactg gaggtgctag cagtgaggac 3120
 agcctgatga ccagcttcct gccaggccct aagcctggag ctccttccc taatggacac 3180
 gtgggtgytg gaggcagtgg cctgctccca cctccaccg cgtctgctg ggctctgcc 3240
 tgtgatgtct ccgtacgtgt ggtggtgggt gagccaccg aggccagggt ggttccgggc 3300
 cggggcatct gcctggacct cgccatcctg gatagtgcct tctgctgtc ccagggtggcc 3360
 ccatccctgt ttatgggctc cattgtccag ctccagcagt ctgtcactgc ctatatggtg 3420
 tctgccgcag gcctgggtct ggtcgccatt tactttgcta cacaggtagt atttgacaag 3480
 agcgacttgg ccaaatactc agcgtagaaa acttccagca cattgggggtg gagggcctgc 3540
 ctactgggt cccagctccc tgctcctgtt agcccatagg ggtgcccgg ctggccgcca 3600
 gtttctgttg ctgccaaagt aatgtggctc tctgctgcca ccctgtgctg ctgagggtgcg 3660
 tagctgcaca gctgggggct ggggcgtccc tctcctctct cccagctctc tagggctgcc 3720
 tgactggagg ccttccaagg ggggttcagt ctggacttat acaggagggc cagaagggtc 3780
 ccatgcactg gaatgcgggg actctgcagg tggattaccc aggtcagggt ttaacagcta 3840
 gcctcctagt tgagacacac cttagagaagg gtttttggga gctgaataaa ctcagtcacc 3900
 tgggtttccca tctctaagcc ccttaacctg cagcttctgt taatgtagct cttgcatggg 3960
 agtttctagg atgaaacact ccacatggg atttgaacat atgaaagtta tttgtagggg 4020
 aagagtcctg aggggcaaca cacaagaacc aggtccctc agccacagc actgtctttt 4080
 tgctgatcca cccctctctt accttttatc aggatgtggc ctggttggtc ttctgttgcc 4140
 atcacagaga cacaggcatt taaatattta acttatttat ttaacaaagt agaagggaat 4200
 ccattgctag cttttctgtg ttggtgtcta atatttgggt aggggtgggg atcccaaca 4260
 atcaggctcc ctgagatagc tggctcattg gctgatcatt gccagaatct tcttctcctg 4320
 gggctctggc ccccaaatg cctaaccag gaccttgga attctactca tcccaaatga 4380
 taattccaaa tgctgttacc caaggttagg gtgttggaagg aaggtagagg gtggggcttc 4440
 aggtctcaac ggcttcctta accacccctc ttctcttggt ccagcctggt tccccccact 4500
 tccactcccc tctactctct ctaggactgg gctgatgaag gcaactgcca aaatttcccc 4560
 tcccccaac tttccctac ccccaacttt cccaccagc tccacaacct tgtttggagc 4620
 tactgcagga ccagaagcac aaagtgcggt tcccaagcc tttgtccatc tcagcccca 4680
 gagtatatct gtgcttgggg aatctcacac agaaactcag gagcaccctc tgcctgagct 4740
 aaggagggtc ttatctctca ggggggggtt aagtgcggt tgcaataatg tcgtcttatt 4800
 tatttagcgg ggtgaatatt ttatactgta agtgagcaat cagagtataa tgtttatggt 4860
 gacaaaatta aaggctttct tatatgttta aaaa 4894

<210> 703

<211> 2904

<212> DNA

<213> Homo sapiens

<400> 703

gtctatgcct tcatgatcag tcttgggggc tgcctgggct acctcctgcc tgccattgac 60
 tgggacacca gtgccctggc cccctacctg ggcaccagc aggagtgctt ctttggcctg 120
 ctaccctca tcttccctcac ctgcgtagca gccacactgc tgggtggctga ggaggcagcg 180
 ctgggccccca ccgagccagc agaagggtctg tcggccccct ccttgctgcc ccactgctgt 240
 ccatgccggg cccgcttggc tttccggaac ctgggcgccc tgcttccccg gctgcaccag 300
 ctgtgctgcc gcatgccccg caccctgcgc cggctcttcg tggctgagct gtgcagctgg 360
 atggcactca tgaccttcac gctgttttac acggatttcg tgggcgaggg gctgtaccag 420
 ggctgcccc gagctgagcc gggcaccag gcccggagac actatgatga aggaaggcct 480
 ctggctgctc taggagtctg atcagagtcg ttgccccagt ttgacagaag gaaaggcgga 540
 gcttattcaa agtctagagg gagtggagga gttaaggctg gatttcagat ctgcctggtt 600
 ccagccgagc tgtgccctct gctcccccaa cgactttcca aataatctca ccagcgcctt 660
 ccagctcagg cgtcctagaa gcgctctgaa gcctatggcc agctgtcttt gtgttccctc 720

```

tcacccgcct gtcctcacag ctgagactcc caggaaacct tcagactacc ttcctctgcc 780
ttcagcaagg ggcgttgccc acattctctg agggcggtcg gatgggcagc ctggggctgt 840
tcctgcagtg cgccatctcc ctggtcttct ctctggatcat ggaccggctg gtgcagcgat 900
tcggcactcg agcagtcctat ttggccagtg tggcagcttt ccctgtggct gccggtgcc 960
catgcctgtc ccacagtgtg gccgtggtga cagcttcagc cgcctcacc gggttcacct 1020
tctcagccct gcagatcctg ccctacacac tggcctccct ctaccaccgg gagaagcagg 1080
tgttccctgcc caaataaccga ggggacactg gaggtgctag cagtgaggac agcctgatga 1140
ccagcttcct gccaggccct aagcctggag ctcccttccc taatggacac gtgggtgctg 1200
gaggcagtg cctgtctcca cctccaccog cgtctgcggg ggcctctgcc tgtgatgtct 1260
ccgtacgtgt ggtggtgggt gagcccacog agggcagggg ggttcggggc cggggcatct 1320
gcctggacct cgccatcctg gatagtgcct tctgtctgtc ccaggtggcc ccatccctgt 1380
ttatgggctc cattgtccag ctccagcagt ctgtcactgc ctatatggtg tctgcccgag 1440
gcctgggtct ggtcgccatt tactttgcta cacaggtagt atttgacaag agcgacttgg 1500
ccaaatactc agcgtagaaa acttccagca cattgggggtg gagggcctgc ctactgggt 1560
cccagctccc cgctcctgtt agccccatgg ggtgcgggg ctggccgcca gtttctgttg 1620
ctgccaaagt aatgtggctc tctgtgcca ccctgtgctg ctgaggtgctg tagctgcaca 1680
gctgggggct ggggcgtccc tctcctctct cccagctctc tagggctgcc tgactggagg 1740
ccttccaagg gggtttcagt ctggacttat acagggaggc cagaagggt ccatgcactg 1800
gaatgcgggg actctgcagg tggattacc aggtcagggt ttaacagcta gcctcctagt 1860
tgagacacac ctagagaagg gtttttggga gctgaataaa ctcagtcacc tggtttccca 1920
tctctaagcc ccttaacctg cagcttcgtt taatgtagct cttgcatggg agtttctagg 1980
atgaaacact cctccatggg atttgaacat atgaaagtta tttgtagggg aagagtcctg 2040
aggggcaaca cacaagaacc aggtccctc agcccacagc actgtctttt tgctgatcca 2100
ccccctctt accttttatc aggatgtggc ctggttggtc ttctgttgcc atcacagaga 2160
cacaggcatt taaatattta acttatttat ttaacaaagt agaagggaat ccattgctag 2220
cttttctgtg ttggtgtcta atatttgggt aggggtgggg atccccaaca atcaggtccc 2280
ctgagatagc tggtcatttg gctgatcatt gccagaatct tcttctcctg gggctctggcc 2340
ccccaaaatg cctaaccag gaccttgga attctactca tcccaaatga taattccaaa 2400
tgctgttacc caaggttagg gtgttgaagg aaggtagagg gtggggcttc aggtctcaac 2460
ggcttcccta accaccctc ttctcttggc ccagcctggt tccccccact tccactcccc 2520
tctactctct ctaggactgg gctgatgaag gactgcccc aaatttcccc taccaccaac 2580
tttcccctac cccaacttt ccccaccagc tcacaaccc tgtttggagc tactgcagga 2640
ccagaagcac aaagtgcggg tccccaaagc tttgtccatc tcagccccca gagtatatct 2700
gtgcttgggg aatctcacac agaaactcag gagcaccccc tgcctgagct aaggagggtc 2760
ttatctctca gggggggttt aagtgcggt tgcaataatg tcgtcttatt tatttagcgg 2820
ggtgaatatt ttatactgta agtgagcaat cagagtataa tgtttatggt gacaaaatta 2880
aaggctttct tatatgttta aaaa 2904

```

<210> 704

<211> 4034

<212> DNA

<213> Homo sapiens

<400> 704

```

aaccagcctg cagcgctgg ctccgggtga cagccgcgcg cctcggccag gatctgagtg 60
atgagacgtg tccccactga ggtgccccac agcagcaggt gttgagcatg ggctgagaag 120
ctggaccggc accaaagggc tggcagaaat gggcgccctg ctgattccta ggcagtggc 180
ggcagcaagg aggagaggcc gcagcttctg gagcagagcc gagacgaagc agttctggag 240
tgctgaacg gccccctgag ccctaccgc ctggccact atggtccaga ggctgtgggt 300
gagccgcctg ctgcggcacc ggaaagccca gctcttgctg gtcaacctgc taacctttgg 360

```



```

tggtcattgg gctgatcatt gccagaatct tcttctcctg gggctctggcc ccccaaatg 3480
cctaaccag gaccttggaa attctactca tcccaaatga taattccaaa tgctgttacc 3540
caaggtagg gtgttgaagg aaggtagagg gtggggcttc aggtctcaac ggcttcccta 3600
accaccctc ttctcttggc ccagcctggg tccccccact tccactcccc tctactctct 3660
ctaggactgg gctgatgaag gcaactgccc aaatttcccc tcccccaac tttccctac 3720
ccccaacttt cccaccagc tccacaacct tgtttggagc tactgcagga ccagaagcac 3780
aaagtgcggt ttcccaagcc tttgtccatc tcagcccccga gagtatatct gtgcttgggg 3840
aatctcacac agaaactcag gagcaccccc tgccctgagct aaggagggtc ttatctctca 3900
gggggggttt aagtgcggtt tgcaataatg tcgtcttatt tatttagcgg ggtgaatatt 3960
ttatactgta agtgagcaat cagagtataa tgtttatggg gacaaaatta aaggctttct 4020
tatatgttta aaaa 4034

```

<210> 705

<211> 6976

<212> DNA

<213> Homo sapiens

<400> 705

```

gaagctggac cggcaccaaa gggctggcag aaatgggcgc ctggctgatt cctaggcagt 60
tggcggcagc aaggaggaga ggccgcagct tctggagcag agccgagacg aagcagttct 120
ggagtgcctg aacggccccc tgagccctac ccgcctggcc cactatgggc cagaggctgt 180
gggtgagcgg cctgctgcgg caccggaaag ccagctctt gctgggtcaac ctgctaacct 240
ttggcctgga ggtgtgtttg gccgcaggca tcacctatgt gccgcctctg ctgctggaag 300
tyggggtaga ggagaagttc atgacctagg tgctgggtga gtcactacat cctccttct 360
tcctgttcca gatacatgcc acctggcatg tgggacagga gtacctctgc cctgggagct 420
gcttgaggag agaggtgggtc tgctgggaag gcattgctgg gcaggagggt gacctgggc 480
tgagggggca caccaagaga aagaagagaa taccaaggac ataccacagt cacctctgga 540
tcctgtgtcc tgcacagagc ctggctcata ggagacactg gagaaatgct cctaaccttt 600
ggctagccct ttataaattt atagcgatta tctcatthaa tgcttacaac caccatttga 660
ggtgatccat ttacagaga aggaagcaga ggcttttaag aggttaggta agtcttagcc 720
aaagccaaat agcagctgaa cagtagagct gggactccat caaggtctcc cagccggagc 780
ttgctcctac ccctaggaca aggggtggac tcctgactct gcagataaat tctacaaaag 840
ccacagaagg caagtagtaa ccattgtgtg acaaccctc accccagga agagggggcc 900
ctgtgaggat tgcaggctct ggagtcacac tgcttgttga aacgctgcct cttaccctcc 960
ctaggctctgc gcctttgaat aagtatcact tmttagttgc tccatgcctc agtttgtcca 1020
tctgaaaatg ggggcatctg taatgcctgt gttatgagga gtaaattaca gcatccctgt 1080
gaagacgtag cacagtgtcg agtacggaat gttatttcca tccttctcac ggagcttgg 1140
tccccttccc cttgcccttt acttgtccca gccattgact catactactt cccttcttgc 1200
aggcattggg ccagtgtctg gcctgggtctg tgtcccgctc ctaggctcag ccagtacca 1260
ctggcgtaga cgctatggcc gccgcgggcc ctcatctgg gcactgtcct tgggcatcct 1320
gctgagcctc tttctcatcc caagggcgg ctggctagca gggtgctgt gcccgatcc 1380
caggccccct gagctggcac tgetcatcct gggcggtggg ctgctggact tctgtggcca 1440
ggtgtgcttc actccactgg aggcctgct ctctgacctc ttccgggacc cggaccactg 1500
tcgccaggcc tactctgtct atgccttcat gatcagctct gggggctgcc tgggctacct 1560
cctgcctgcc attgactggg acaccagtgc cctggccccc tacctgggca cccaggagga 1620
gtgcctcttt ggctgtctca ccctcatctt cctcacctgc gtagcagcca cactgctggt 1680
ggctgaggag gcagcgctgg gccccaccga gccagcagaa gggctgtcgg cccctcctt 1740
gtcgcccccac tgctgtccat gccggggccc cttggctttc cggaacctgg gcgcctgct 1800
tccccggctg caccagctgt gctgcgcgat gcccgccacc ctgcgcgggc tcttcgtggc 1860
tgagctgtgc agctggatgg cactcatgac cttcacgctg ttttacacgg atttcgtggg 1920

```


Leu Ser His Ser Val Ala Val Val Thr Ala Ser Ala Ala Leu Thr Gly
50 55 60

Phe Thr Phe Ser Ala Leu Gln Ile Leu Pro Tyr Thr Leu Ala Ser Leu
65 70 75 80

Tyr His Arg Glu Lys Gln Val Leu Ile Gly Gln Trp Val Glu Ser Gly
85 90 95

Trp Glu Gly Trp Ser Gly Phe Leu Gly Gly Gln Leu Ala Gln Asn Leu
100 105 110

Val Ser Gly Lys Gln Leu Trp Arg Met Leu Leu
115 120

<210> 707

<211> 150

<212> PRT

<213> Homo sapiens

<400> 707

Met Val Gln Arg Leu Trp Val Ser Arg Leu Leu Arg His Arg Lys Ala
5 10 15

Gln Leu Leu Leu Val Asn Leu Leu Thr Phe Gly Leu Glu Val Cys Leu
20 25 30

Ala Ala Gly Ile Thr Tyr Val Pro Pro Leu Leu Leu Glu Val Gly Val
35 40 45

Glu Glu Lys Phe Met Thr Met Val Leu Gly Glu Ser Leu His Pro Pro
50 55 60

Ser Phe Leu Phe Gln Ile His Ala Thr Trp His Val Gly Gln Glu Tyr
65 70 75 80

Leu Cys Pro Gly Ser Cys Leu Glu Gly Glu Val Val Cys Trp Glu Gly
85 90 95

Ile Ala Gly Gln Glu Gly Asp Pro Gly Leu Arg Gly His Thr Lys Arg
100 105 110

Lys Lys Arg Ile Pro Arg Thr Tyr Pro Ser His Leu Trp Ile Pro Gly
115 120 125

Pro Ala Gln Ser Leu Ala His Arg Arg His Trp Arg Asn Ala Pro Asn
130 135 140

Leu Trp Leu Ala Leu Leu
145 150

002230-682990

<210> 708

<211> 371

<212> PRT

<213> Homo sapiens

<400> 708

Met Leu Phe Pro Ser Phe Ser Arg Ser Leu Val Pro Leu Pro Leu Ala
 5 10 15

Leu Tyr Leu Ser Gln Pro Leu Thr His Thr Thr Ser Leu Leu Ala Gly
 20 25 30

Ile Gly Pro Val Leu Gly Leu Val Cys Val Pro Leu Leu Gly Ser Ala
 35 40 45

Ser Asp His Trp Arg Gly Arg Tyr Gly Arg Arg Arg Pro Phe Ile Trp
 50 55 60

Ala Leu Ser Leu Gly Ile Leu Leu Ser Leu Phe Leu Ile Pro Arg Ala
 65 70 75 80

Gly Trp Leu Ala Gly Leu Leu Cys Pro Asp Pro Arg Pro Leu Glu Leu
 85 90 95

Ala Leu Leu Ile Leu Gly Val Gly Leu Leu Asp Phe Cys Gly Gln Val
 100 105 110

Cys Phe Thr Pro Leu Glu Ala Leu Leu Ser Asp Leu Phe Arg Asp Pro
 115 120 125

Asp His Cys Arg Gln Ala Tyr Ser Val Tyr Ala Phe Met Ile Ser Leu
 130 135 140

Gly Gly Cys Leu Gly Tyr Leu Leu Pro Ala Ile Asp Trp Asp Thr Ser
 145 150 155 160

Ala Leu Ala Pro Tyr Leu Gly Thr Gln Glu Glu Cys Leu Phe Gly Leu
 165 170 175

Leu Thr Leu Ile Phe Leu Thr Cys Val Ala Ala Thr Leu Leu Val Ala
 180 185 190

Glu Glu Ala Ala Leu Gly Pro Thr Glu Pro Ala Glu Gly Leu Ser Ala
 195 200 205

Pro Ser Leu Ser Pro His Cys Cys Pro Cys Arg Ala Arg Leu Ala Phe
 210 215 220

Arg Asn Leu Gly Ala Leu Leu Pro Arg Leu His Gln Leu Cys Cys Arg

002690-034990


```
<210> 709
<211> 141
<212> DNA
<213> Homo sapiens
```

```
<400> 709
tacggcgtgg  tggggagggc  ggtaccccac  aaataacacn  nacaccccat  cctatctgtg  60
tccacanata  aantgactca  ttctctctct  cgcatanccc  actntccccc  ngcgataaccg  120
taacnaance  ctccccctt  t
                                                    141
```

<220>

<400> 713
 nntgggtcgcc tgnngcgtnta ctctaaagga tntactatnc atatggantc naanacgact 60
 cactacacgg cncctcncgg agccnnggtc agtgcctnct nggagacctt ctctggggca 120
 ggangagcac tnggtatggt caggtatcnc ttentaaana tacnncctc cg 172

<210> 714
 <211> 112
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(714)
 <223> n=A,T,C or G

<400> 714
 nttgcgtgcc tggacgtnta ctctgcanga tctactactc atgngaattc taantacgga 60
 ctcactatnc ggcancgcag ggcgcagcagg gaangggcca cctcccagtc tc 112

<210> 715
 <211> 326
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(326)
 <223> n=A,T,C or G

<400> 715
 tactctanag gatctncgng tcatntggat tctatntcga ctcactctag ggctcnagcn 60
 gtcngccggg caagttattc ggatcgtcgg gntccgagct tcgcaattaa ntgtgccatc 120
 gttctncaac gttcctgact nggaancccc ngcngttcng atccncnggt acctagctcc 180
 anntccccg tntccttctt ggngtntcat naangaggac cncctctgat cnccttctc 240
 taatctgcnc acnctgaacg nccaatggac atngtgcgtt taatntanna ggcccgnctc 300
 gngtgccctt cccgtnannt cagctc 326

<210> 716
 <211> 122
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(122)
 <223> n=A,T,C or G

<400> 716
 nntgcgtcgc ctgngcgtnt actctagatg atctgantag tcatatggat tctaatacga 60
 ctcannatag ggctctagcg nggatncnga ttcgtcntcc ngattcantg acnccggtan 120
 ca 122

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(131)

<223> n=A,T,C or G

<400> 720

```
tccatcctaa tacgactcac tatagggctg ccaacctgcc atccactact gaggaagacc 60
cgnaactta ggggctcact gcgagccacc ggccacaggt cgtatagggc aaagcacgng 120
gaagaccccc t                                     131
```

<210> 721

<211> 121

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(121)

<223> n=A,T,C or G

<400> 721

```
tccatcctaa tacgactcac tatagggcgg ntgantnctg gcgaaaggct tacaattaag 60
naggaaaaan ganccaacaa ctaaaaaaaa nncggncgtg ncagcttnga tgactngtcc 120
a                                     121
```

<210> 722

<211> 246

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(246)

<223> n=A,T,C or G

<400> 722

```
anctggagtc gcgcgctgca gtcacattgt ggatccanaa aatcggcaca agctctcntg 60
gnttcntcga tatgaanaac actaatccca tgtngtntgn gtctccgtga ttcattccctc 120
gcacnggtcc ccntccnaac cnttgcatag gtgttatggt gtantctccc cagtgcacaa 180
agattnacac tctctcantg tctganatat gcacgagttc attgtcctgt cnccgtnaac 240
atcaag                                     246
```

<210> 723

<211> 160

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature


```

<400> 726
ttgggtgggt tgggtggggg naaatttncc catttggtg ggtttggggg ggnaaatact 60
tccgccttt tnggtnccca aaganacnaa gggggagtcc cttnatagag gnagncgat 120
ncntcncaac nacntngact ttgnccatgg ggagnaaggt gg 162

```

```

<210> 727
<211> 120
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(120)
<223> n=A,T,C or G

```

```

<400> 727
gtgtgggtgg ggaattccat tgtggttggg ggnaaatctc cgcttgtcca aagnacaggg 60
gggtcncctt anagngnagg gggttcctcc ccaccacttg ncttgnccat tngagnaag 120

```

```

<210> 728
<211> 130
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(130)
<223> n=A,T,C or G

```

```

<400> 728
gaccactgc agcgttnaac ttagcttgga ccgagctcgg atccctagtc cgtgtggtgg 60
aattccatgt gtcgagagag gggcaaatac nctccaanac ancncctca tgctcnacac 120
atattcgcat 130

```

```

<210> 729
<211> 182
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(182)
<223> n=A,T,C or G

```

```

<400> 729
cngactgctn gcgtttaaac ttaagcnagg taccgaacgg ggatnnacga ctantgatcg 60
gctggctgct tccagtcgat tanatttggtg aaaaagctga accncngccn gttaaggggg 120
annatgcaaa anatncatcc nnetgccccn taaactgntc tntccnaggg aaaaaangga 180
ag 182

```

```

<210> 730

```

<211> 678
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(678)
 <223> n=A,T,C or G

<400> 730
 cactcncact ccggacctag gcncttcacc actgctctct tctcctcct cctcctcntc 60
 ctcggggctg ggggaccttc ccagtgacc atctcacttt ggctgaancc cactcggggc 120
 agcctgagtt tggggctctt ggccttctca cctcctcgg cccctcctt ggcccgacc 180
 aggccaaacc ggggcagccg taccttgagc ttgtgtcgg cctctcctc cccctctgcc 240
 acctggtact cggcatggtt gcccccgga tggogagagc tccacgtcgg gcagtgagaa 300
 gcagaaagta cgctcggccc ctgggggctg ctctcagca ccctcgcccc ccaccctagc 360
 tctggcccc agtgtgggca acttcagcct cagcccaccc tcgcctgtgg ccgcctcgcc 420
 cgctgtgcc tctcggctta gccccacgtc caactcaagc tggggcactg tcacggtggg 480
 catcttaaag acaccctcac ccaccagcag ctaccacct gcaacctggg ctccaggcaa 540
 aaaaagggtc acctggggca nctgaacct gtacctgctg tgccctctgc tgaangaat 600
 gttatctgaa cctgctgccc tgggggtact gccttcccaa aaccgggtca antccacctg 660
 ttggaaggna aatncccc 678

<210> 731
 <211> 135
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(135)
 <223> n=A,T,C or G

<400> 731
 gagatccgac gtcacccccct tccggcgggc caagacgctg caactcccga ggcngcccaa 60
 atatcttttg aagagcgctc ccagcccaac acaatggaat tccaccacac tggnttagtg 120
 gatccgagct aagcc 135

<210> 732
 <211> 660
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(660)
 <223> n=A,T,C or G

<400> 732
 gcttggtacc gagctnggat ccctagtaac ggccgccagt gtgctggaat teggtttct 60
 tcaatcagnt nacgagctgc atggtctgct aacattgtca taattgctgg catagattac 120


```

tgaaaataaa gaaaaaaaat tgaagctgcc tatcaagttt tggattattc aaaaacttcc 180
tacaagttat tttacttcaa ccatgttatt acaaataatt taatgaatac tttagagact 240
ttaattacaa aaaactgaga tagtaaaagc aagtaataaa agctgaaatt acttagctat 300
ttgataatta cataaattat tatggtccat tcaacttttc tagtgtttag tttatacacc 360
aggaagaactt tctatttcta ctaacattta taaagtatgc taacctatta tttaaacgca 420
tccactatta ggatttttatg gcctaaaacg tgatacagtt cagtatcttg atgtcaaaac 480
ttttaagca agtagggatt aagttcaagt gaatgtgatt ttctttcttc ccagtagggg 540
cttctgaata actcagnaag gctcacttcc attatcttac tttataaaaa aatgctataa 600
gacagaatgg gccgacgtgg nggctccacc tgtatccacc tttggaggcg agnggcgaat 660

```

<210> 733

<211> 836

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(836)

<223> n=A,T,C or G

<400> 733

```

aattaatgac tttttttccg ccctgccaaag ctagtttgtc taaatataat gtaaagaaat 60
tagctactca ttttctgggc cacgaagggt cctaaaatgg gaagaagtgg agatctgacc 120
ttgtagtttc taaatacact aaactgggag tgccatggat ggctttcagg atgtcctgaa 180
tcccctataa ttgtatacaa aatcgtgagt ttttaaaaaac tgggttagag ctattgggtc 240
ctcagagtct caggcatctt agaccccaa aaagggttaag gactactgac ttaaccaatt 300
aggtttgagt ggcattggct ttgaagaaaa gcagaggaaa gatataattt ataattctgg 360
gcaacaaaaa agtggatgtg tgccagcatc ttagagtaga atcctcttaa aaggatagca 420
ctgcatatga actagtaggt ttaaccaggt gcataattag gcgaagtagc tcatttttct 480
gttagaattc ttttttattt gggaatgggc aagcttttac agcttttacc ttgccaatga 540
atacctggaa tttaaaaaat ctgttagggc atattgcca taaagttttt tttcctagat 600
catatattca gtaaataatg ttgtagcttt atttcaatcc cccaattcat tgagggttga 660
aacaatttga atggtttgag tgtagaagct aagttatttc tgtagaggct aagggcattt 720
ataccaanat atgttagact tngngntcct gttaaccatg ctgtanacaa taggaattac 780
tgtatatcca cattttaatt ttaacatctt ctgctttgnt gntggtttga gangga 836

```

<210> 734

<211> 694

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(694)

<223> n=A,T,C or G

<400> 734

```

nagtntctatt tncactaaac tngngagtgc ttggatggct ttcaggatgt cctgaatcct 60
ctataattgt atacaaaatc gtgagttttt aaaaactggg ttagagctat tggttcctca 120
gagtctcagg catcttagac ccccaaaaag gttaaggact actgacttaa ccaattaggt 180
ttgagtggca ttggctttga agaaaagcag aggaaagata tattttataa ttctgggcaa 240

```

```

caaaaaagtg gatgtgtgcc agcatcttag agtagaatcc tcttaaaagg atagcactgc 300
atatgaacta gtaggtttta accagtgcac atttaggcga agtagctcat tttctgttta 360
gaattctttt ttatttgga atgggcaagc ttttacagct tttaccttgc caatgaatac 420
ctggaattta aaaaatcttg ttaggcataat tgcccataaa gttttttttc ctagatcata 480
tattcagtaa atatgtttgt agctttatct caatcccca attcattgag gggtgaaaca 540
atttgaatgg tttgagtgt gaagctaagt ttttctgtga gaggctaagg gcatttatac 600
caagatatgt tagacttgtg gttcctgtta accattgctg tagacaatag gaattactgt 660
atatccacat ttttaatttt aacatcattc tgtc 694

```

<210> 735

<211> 126

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(126)

<223> n=A,T,C or G

<400> 735

```

ncnttgaaac ngggtgacca gacttcaggc ctgtgcgctc aatcgtggag aatctcgtgc 60
cgaattcggc acgagctctc ctctctctct ctctctctct ctctctctct ntctctctct 120
ctctct 126

```

<210> 736

<211> 165

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(165)

<223> n=A,T,C or G

<400> 736

```

cagaagcctt taaaccggtt ngaccagact tcaggcctgt gcgctcaatc gtggagaatc 60
tcgtgccgaa ttcggcacga gtctctctct ctctctctct ctctctctct ctctctctct 120
ctctctctct ctctctctct ctctctctct ctctctctct ctctc 165

```

<210> 737

<211> 125

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(125)

<223> n=A,T,C or G

<400> 737

```

ggnagccctt ttaaccgttt gtccagactt caggcctgtg cgctcaatcg tggagaatct 60
cgtgccgaat tcggcacgag tctctctctc tctctctctc tctctctctc tctctntctc 120
tctct 125

```

```

<210> 738
<211> 137
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(137)
<223> n=A,T,C or G

```

```

<400> 738
ggagncnctt gancaggatg accgacttca ggccctgtgcg ctcaatcgtg gagaatctcg 60
tgccgaattc ggcacgagtc tctctctctc tctctctctc tctctctctc tctctctctc 120
tctctctctc tctctct 137

```

```

<210> 739
<211> 970
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> misc_feature
<222> (1)...(970)
<223> n=A,T,C or G

```

```

<400> 739
aggcctattht aggtgacact atagaacaag tttgtacaaa aaagcaggct ggtaccggct 60
cggaattcgc ggccgcgtcg acggcccttn gtgccactag ntctttcatt cttccccccc 120
atcaatcagt gaacttttta gcctactcaa agctttgctc caatgcatag gatttatgat 180
tgtggggatt tccagataat ataaatattc aacatgaata ttttaaatta aggcatgaga 240
catttttctt aactgagcat agccatgaac ctctcacgct tgttcctctg tgtcagtttg 300
tancactgaa tacagcagcc ctccataaaag tccaggcagt gcacaggctct tgacatgatg 360
aagtgcagtg ttgctatggg gatthttgcag ctggccaaat agtcactggg tgattttacc 420
cagcaggaga tttttgcaaa aatttctctg gtgagagtga aatcaaactc ctattttgnt 480
tctctctctc aagctgnagt taagatggat taatgagtag ttttagatta attaactctg 540
aagagaaaaat gggagaaaaa tgaggaaggg tgttggcaga agtcattgct ggaatccttc 600
tgaaggaggat actgacttca cttgcaaaga cnagagacta naagacaatg aagttaaact 660
tggcctgtct ctcatatgat agatgctgag agtcaggntc agggaaattht aattctgtca 720
tacgcatatn ggattatgtg gtcattggat tgttggcact aaccngcctn taatcagnat 780
aagaaaagtg ttttggtaga naaagaaaaat tatggccag aaaaacctgg aanacttga 840
aaaaatgntn gggggccttg ggtgggtggc tnaaaanacc ccctggggat ntttaaacca 900
aaantgaaga agggaaaaat ntttccctnt nttttntttt tttgccccct tgggatttgg 960
tttnttttcc 970

```

```

<210> 740
<211> 739
<212> DNA

```



```

ggaaaaaaat tngnnngggg gccnttttggg tggggggggt tnaaaaaacc ccctnngggg 1080
ttttttaagc ccaaaagggg gggaggggna aaanggtnc cttntttttt ttttnngccc 1140
cccttgggga atggnnttant tcanggggcc c 1171

```

<210> 742

<211> 739

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(739)

<223> n=A,T,C or G

<400> 742

```

gntgtcnaaa aagcaggctg gtaccgggtcc ggaattcgcg gccgcgtcga cggcccttgg 60
tgccactagt tctttcattc ttcccncca tcaatcagtg aacttttttag cctactcaaa 120
gctttgctcc aatgcatagg atttatgatt gtggggattt ccagataata taaatattca 180
acatgaatat tttaaattaa ggcagtagac atttttccta actgagcata gccatgaacc 240
tctcacgtct gttcctctgt gncagtttgt agcactgaat acagcagccc tcctaaaagt 300
ccaggcagtg cacaggctct gacatgatga agtgacgtgt tgctatgggtg attttgcagc 360
tggccaaata gtcactgggt gattttaccc agcaggagat ttttgcaaaa atttcctggg 420
tgagagtga atcaaaactcc tattttgttt ctctctgca agctgnagtt aanatggatt 480
aatgagtact tttagattaa ttaactctga agagaaaatg ggagaaaagn gaggaagggt 540
gttggcagaa gtcattgctg gaatccttct gaagggagta ctgacttcac ttgcaaagac 600
aagagactan aagacaatga agttaaactt ggctgtctn tcatatgata gatgcttgag 660
agtacaggnt cagggaat ttaattctgn catacgcata ttggattatg tgggtcatgg 720
ctttgtttgg cncctaacc 739

```

<210> 743

<211> 610

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(610)

<223> n=A,T,C or G

<400> 743

```

ctgtccttat ttcttttagca aaaatttccc aagagaagaa ttgctgggat aatgcacatt 60
taaatttttg atagacattc ccaaataatta tacctgtttt tgagaccttt aattcctgtt 120
gtcaaattgc cctatatatg gagtaataaaa caggatttaa agaaatgagg actaaaaaaa 180
gattatatat aacccaacat aaaggcaacc tcttaggcgt tgacagaaac tgacaacttt 240
ttatctgtgg gtgcgatcca ttataagtaa cctgagcacc ttattttttc tttttaaact 300
ctaggtagga taccggagggt ccacaaattt ttcataagaa atattttttc tctgccttat 360
gagattttta aaaatattat actgcttcaa ttgcatcaaa agaaatggac cctaatatct 420
atgatgaagg atttggaggt agaagacctg agtttcaatt ttggcatggc tgtttgtcta 480
gctctngat cttggacagg tcaattgact tggtttaatc ttctcatcca tttagnngag 540
acagcaccac tattcacagg actattgn cn gaattaccag acaatagcat agngaaaaat 600
ataangcctt 610

```

<210> 744
 <211> 127
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(127)
 <223> n=A,T,C or G

<400> 744
 ttnacctccc tggaccgggc ccccttccc cgggcggntc ccccgggctg caggaattct 60
 gcacgagga gagagagtn gagagagaga gagagagaga gagagagaga gagananaga 120
 gagagag 127

<210> 745
 <211> 458
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(458)
 <223> n=A,T,C or G

<400> 745
 gatatcccgg gattcgcggc cgcgtcgacg tggcctctag tttgtcctgg tccaaagcag 60
 ggaagctggg ctacgtcctg cccaggtcag ccttaggtta agggctgcct gggggagga 120
 acttcctggg ccttcgggtc tctgtgcaact ggggtggctc ctgtggcca gaatgccctg 180
 gagaagggtc ctactggaag cgaagggtgca gggcagcagg gcctgaggcg caggagctgg 240
 tggaggctcc cagcacaggc cgcgcggcca gtcacatcac tgctgatggg ggggggactt 300
 ggggagtttc ccccagaaat gggagggtctc acagtccccg tgctgcaatg ctgtcgggtgc 360
 actgngneng caatgtgctc atggncaactt gctttttctc tgtggccccg gccgatttat 420
 ccagcanngc accctcttc tncctctccg anaaagcc 458

<210> 746
 <211> 893
 <212> DNA
 <213> Homo sapiens

<220>
 <221> misc_feature
 <222> (1)...(893)
 <223> n=A,T,C or G

<400> 746
 aagcaggctg gtaccgggtcc ggaattcgcg gccgcgtcga cgtggggagt tagctctctg 60
 gaccccgctca tagagtaagt catcgataga gcatttgctt gatggggact tccagaaggc 120
 canngaaaagt cctgccgact tcctggggaa gcccatccgc acgtgggggtg aggggtcccca 180
 natggaagca gctgtgtatg cagggagggg gcagaggctg ctgccaatgg gcatgtccct 240

```

tacctgaaag ggccacctct ccaggtgaca tgtcctgggg gagccggggc cgtctgctcc 300
ggccagaggc gctcagctca ggccacacca ggcagggcac ctcccaacct ggacaggtgg 360
ggaccaaggt ggccttggac aaaactctct gtgtttgcca agcacccaat cggacacaga 420
gagtcaacca caccacagtc acatggtgtc cacacngcag ggggtcaagga ggcccggccc 480
ctccccctca gacgtccctg ggctctctgg agtcagcaag gacgaggacg gcattgcctt 540
tcgagacagg aaggaggatga cctcctcccg gcggcatcca ggctcngctt ctccggagag 600
gagagggggc tacttgctgg ataaancggc cggggccaca gagaaaaagc aagggtgacca 660
tgagcacctt gcaaacacag tgcaccacc agcatttnag caccnnggac tgtgaagacc 720
tcccatttct tcggggggaa acncgcccac ngttcccccc acctcacta gtgnattgtg 780
acctgggggn cggggcgacc cctgtngctt gggnnagecc tcnccccagg tttctnnggc 840
ngcccnttaa nggncctng nttggcccct tggecncett tncgcttttc cca 893

```

<210> 747

<211> 738

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1) ... (738)

<223> n=A,T,C or G

<400> 747

```

gatatcccg gaattcgcg ccgcgtcnac gaagcacaga cctgngccct gctctcatgg 60
ggcagactgc catttgtcat tnattactga aggaaagga tcctcagttt gcttgtggac 120
atttcaaatt tgaggtgaga gttggataag taagaataaa gctgctcttc aaagagatga 180
atatagaaaa agaaacaaga tacagncttg gcagtaaggc tgggaggaag gggaaaaggt 240
aataaagaat gaaagagtga gaaatgtgag caggagctga acacagaaaa gttcagngac 300
agaagcanaa ggagggaaga agggaggagg gtccctttca cagaggctca cgaggatgct 360
ttatgngtgc catgcagtcc atgttcagga tgtctgcttc ttanctctct acttttctaa 420
tanaaatttg gatacttact gatcctacat atgtaacagg gagagaaggt gaatttcaa 480
gcantaaatt gaaaaattgt tcacaatttc atttttttaa aaaaggagc taacagaaga 540
agaggttaat gtggttaatta taggatgnct cttgcgacac atgaatgnat ctggtatcat 600
ctgagtggga ggggagctgt ctctctgacc caaaaggatc ctttcgttan ccngnactta 660
ngtcccaaaa cctcaccacc ttggagaaat natttccttt tgggggtntc attaaancct 720
tttgncccc gcaaaagc 738

```

<210> 748

<211> 647

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1) ... (647)

<223> n=A,T,C or G

<400> 748

```

ctntgtggcg gtggctgtct catttgggtg gacttttttg gtcgtaggaa cctggtatng 60
aggctcgagag taagacgggc tattagtagt cgcacggag ttatttgtga aaacctggtt 120
agggcctctg tctccgctgc gctcgcctaa attggtatgg ctcgacttgg aaacacggtt 180

```

```

ctaacacgcg ttgttagcgc ccttgctagc atgtgaagga cactggccct accaagaaag 240
attcgagtcg ctccctcccg tatcgttcac ggaggcgata ttactcttc ttactacggg 300
tacttcgaga ttgtctgtga agtttaagac tactaaaaag agtattaagc ctatcgggaa 360
ttagctagat cgacacgcta aaaccaaggg caatcggcgg aaatatagag gcaccaataa 420
tagggcctac agaaggcccg agggttagac tcacgtttta taccggccac gggagaaata 480
aaaagataaa gtatacatcg tttagcggtc ctcggaagcc ttcggcttta atgccaaagga 540
gtcgggaagca tcgtcggcga gtaataaaact ccatcgccgc gagactatct acgacgccct 600
ccttaanatc cgtaaattac tcccggaaag agtatttagg cggtctct 647

```

<210> 749

<211> 642

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(642)

<223> n=A,T,C or G

<400> 749

```

ctntgtggcg gtggntgtct catttggggtg gactttttgg gtcgtaggaa cctgggtatgc 60
aggtccgcgg agcgtgggct ctcgtcgtgg atgttggggg ttggtgtggt gccggttggt 120
tttggttctg ttgagcgtag tgtgtttgaa ggtagcggt cgtgtcttgc ttgtggtttg 180
gtgtttaggg cgggtgggga gggtgtgtgt tagctgttgt atgtcatatt gttggtgttg 240
ctgccctgtg ctgtttgtcc ttggttattg tggttgttac ccgcctgtg tggaagtgtt 300
gtggcagggc gggaatttaa gtgggagagt tgtgggaccc gtggttgttg ttacgttgct 360
gcttttgtcg tgggcggtgg cggcgcgctc gataattaga attggatacg gagtgtataa 420
tactttctagt aaatggggac ctagtgcctg acttcccggg atagggatct atgcgaagtc 480
cttaggatag tctttgataa gtttaacgcc cagcaccta aaattataca cgattagacg 540
cataacgact cctccaggaa agataaagaa tctcacatat agaacgggac ccatacacg 600
tcggatagga aacaagagaa ctaattttng ttaaaaagac tt 642

```

<210> 750

<211> 639

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(639)

<223> n=A,T,C or G

<400> 750

```

tttgtggcgg tgggtgtctca tttgggtgga tttttgggtc gtaggtaacc tggatatngag 60
gtatagatgc cgattggtcc cgacgagcgt cagcataaat tcggtagttt cgcccttttt 120
agaaggcgct agtactcgga acttcacttc atctcggtag tttacttttg cgtatatagc 180
cttctccctc gaagactagc cgtcacattc gttccctagg aatcgtttct gccctaaga 240
atccgagagc gagatcccgga aactagagga accttagaag agtcgtatct ccacaaggac 300
cccacagtca ttccgggaaa atccctagga ccatacgggt aggattcccc cggaaccccg 360
agcaaaagctc atgatttccc acaccgcgag agcgccata accctatccc atttcttcgg 420
gttatcgagg atattacgat caagccgaga gaaccgctag aaccgctttc ttcgctttct 480

```



```

cacggaacct ataagtagaa agagaaactc aggtcttaag ggggcgcttc ggctaacgaa 540
acttctactt acgaagagag tatctagaca ttaagtcata aaaatccact acgcacctcg 600
tgtacgatat catcgggagc ggttcataga cgggtgtccg 639

```

<210> 751

<211> 637

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(637)

<223> n=A,T,C or G

<400> 751

```

cttttgtggc gnggtgtct catttggtg gatthtttggg tcttaggnaa cctggtatng 60
aggcagctct gagccccccc ccccccccc ccccccccc ccccccccta gnggttggg 120
aanacggtgg atacctaaat cgagtnggtt cattaaaagt agttgattac nccctaaaat 180
aanaanaggg cttcgtcggg anaaatcggg aagganaagt ctttntggca tcataanaat 240
actggctcgg gtcctaanat nttaaggng gtcnccgagg gtnttcatac cgataanaaa 300
cgthtttccta tcggcaacgg gcttacctga gggnggactt ctncggngc ggngattnan 360
acgaanacgt agaggattnc cgntacttnt tganatcacn cgtatcatac ttgtaagcat 420
aatnttcctg aaaagtgtta taanaatacg cncgcattat cgctthtttcg tcctagggat 480
gcttaaatgg cgatactgct atagcgggtg agcgttgggt ctcgagnaan aaagcgtgct 540
ctaattgcgtc taaggnttta agnncgttgg tttaaaaata nccttagaaa cctcgaggcg 600
gatactggtt tntttttaac gaaacaaagc accccnn 637

```

<210> 752

<211> 644

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(644)

<223> n=A,T,C or G

<400> 752

```

tntgtggcgg tgggtgctcat ttgggtggat ttttgggtcg taggaacctg gtatgaggtc 60
ttgcgagttg ttggtgtgtc ctgctgttcg gtggttccct tttgagttga gtttgtcctt 120
tgaggttggt agctgctgtt cgtttgtgtt cgtgtagtgc tttgggttga gaggttatg 180
gtggtgggtta cgggtgtattg tcgcccgtgg tcgcgggggtt ggggtgggtc tcggttttgt 240
ggttcatagt agtcttctgc gttcgggtgg gcgggttttg gtgagtagtt tcgttcttgg 300
atgtcccat gaccgcgcct aatctaagta agggttagta gaaacctct cccgatagac 360
acaaccgtcg tccactaaag acctgcctc tgatthtttaa aaggaccgga aaaacatccc 420
ttcaacggaa aaaacggaaa aaaagtcagc gaattcaaag aagccacggg agagaaaaaa 480
gaactaaagt tagtccgtca ttatatgtct ctcggagga ggaagcggcg gtggcggaaa 540
atgaggcggg aagaaagacg acctctatcg gcggcttang ccctaaaagg gcgatacctt 600
acgggatgat aaggacccta ggacgcctcc ttctcggatc gtcc 644

```

<210> 753

<220>

<221> misc_feature

<222> (1)...(657)

<223> n=A,T,C or G

<400> 759

```
ctttgtggcg gtggtgtctc atttgggtgg actttttggg tcgtaggaac ctggtatnga 60
gggctctata gaaagcctct tgtctttaga tacgggcttt ctggtccttc gttctggaag 120
tgtagtagta ggtactgcgg gaaggcgaag agtcctttca aggacgattt acttaagttg 180
gcttattcta tagttccttc gggacataag gtcggtacga tctatactgc gtgggaagct 240
gataggttgg gacttaaggc gaataagaag gaggcggcgg aggtcgcgat taccgcagag 300
atattattta cggcggccgc gggtagccgc ggtcatgcgg aaattttctg aggttcttgg 360
attcctaaga tcgctcccgt cgagtatact agcgacgaac gtaagagtgc cctcacaaga 420
accggtacaa actcaagaag aagttcccat taagcatcgt aagaaacggg aggacgagga 480
cggtaagaag taatcggaga aaggatccta gtngttacga agaagcatcg ttnagctact 540
ttgcgctacc gtttatattt agacgtgttc cgtccttctc cgtgtttana aaaaagggtt 600
attccgacgg gagacttagg cgaatggagg gttccgcggg tganaatcgg ancgggg 657
```

<210> 760

<211> 644

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(644)

<223> n=A,T,C or G

<400> 760

```
ctttgtggcg gtggtgtctc atttgggtgg actttttggg tcgtaggaac ctggtatgna 60
ggaaaagaag taagcctcga agcctatctc cgaccgtatt tatttcgcag aagacggaac 120
tacggacgtc gttaaccccc agtagcccc gtaagaaagg actaaagcga atggaaaagt 180
cgggaattcc ggcggagggg cggcgattac tgaaaggagt aagagtaaga ctattgcat 240
acttgaggcg ttccctctta aaaggcaccg gaaacactct attaaaaaac acccgaagaa 300
gaacaactca tgcgatcggc cgtgtgcagc cgtcaatagt aaagagagcc atgaaccatg 360
ccatccttag accaattagg atgaagaaga ggaggaagat gaggaccaa ccctaccac 420
tcggaaaacc ccgcacgagc ctccgaacaa aatccgggaa ttaaaacggc ggcccacttc 480
cgactctcg tagcgcgagc cgaatagaaa accggaaact acagctaaag ggtcctttcc 540
ggcctgttat ctaccacccc gcaatccgat cctccccccc cctcgtccaa aaaccctaac 600
ctctgcggca acattagagc agaaggagag ggcgatccct tgan 644
```

<210> 761

<211> 647

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(647)

<223> n=A,T,C or G

<400> 761

```

ctttgtggcg gtggtgtctc atttgggtgg acttttttggg tcgtaggaac ctggtatnga 60
ggcgggtact ctctgggata atcgggtataa gtgttgtaaa attgggggta agagaaagtt 120
tcattataag aagtggaagc acgagccggg gtgttttagtc gttaatatta agaccggttt 180
ttgttgtagt tatatagctt gcgcgtgggg aggcataaag aaacattgcg tttcgaggcc 240
ggatgcgggg aaccctcttc ggggtctaga gcgccgcac tgcaaaataa ggactactga 300
cgccgctcat aacgtactca acaatgagtc ggctgcatt aagatttcgg cgaagaaccg 360
tactgcgtct actgatagta tattgcattg atagcggcat gagctttatc acgtgtcgtt 420
ttcgggttgt aagaaggag ttaagtcgat cttcgaggaa gaagagacc caaataaaaa 480
atgactcaaa aaaacctaga agaaacacga cgaaaggaaa aagaacgtta aaactagtag 540
ctcttcggan gagtagcctt agtagggtaa gtctccgtg cgtactgtcc taaggtttgg 600
atagcgcggt tgaatagacg gtcacgcgtc agaaggtaaa aanccgg 647

```

<210> 762

<211> 628

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(628)

<223> n=A,T,C or G

<400> 762

```

cattgtgttg gggtcactga gccactttt ttccagattt tttgtaaaat tgtttcgcat 60
tgtgttccct ttattcgctt gtattaatat ttgcgtagtg gattaaacaa atacttggtg 120
ttgactgtca gtcttagagg actgactaga agtagttttc atttggggct caggaaatac 180
ctactttata tttctagcta attaggaaag tcatTTTTca gttagggttg tgttttgggt 240
caggcactcg ctagctagat gacctaacat gctacttaat ttctgagtgt ttgtgtccat 300
ccctgtagga ttgttgcggg gttaaatgaa attgtgtata tttgtaaagc atttacctca 360
gtgccagac tgtgacagag tagattatta ggcttgctct tatttctgtg attaaattta 420
gtgtcagatt agcaacctat agctacttct aaagctgctg ctgctttctt tgtttagggg 480
taggaagaaa catgctggac agtttgccaa atgagagtta catgatgtgg cttgtgggaa 540
cattctaact tggaacttgc ccatttccag gactttgngg ttcanagatt tttggggata 600
gatgtaaggg ttaaaaaaaaa cngaaaaac 628

```

<210> 763

<211> 147

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(147)

<223> n=A,T,C or G

<400> 763

```

cattgtgttg gggcagagat aaataattcc tctgaaaagt gttttattgg aatttcaa 60
gaaaagctaa ctggataact tacagcatgt ttctgccaat aatctcttan aacaggcctc 120
ttttttttat gcacaccacc ttenggc 147

```


<221> misc_feature
 <222> (1)...(602)
 <223> n=A,T,C or G

<400> 767

```
nnnttttaaaa nctgtntctcc ccgcggtggc ggccgctcta gaactagtgg atcctttcca 60
cctggtttgt tttcagtgtt taatcctatt agtatcagca ggatataggt caggatatca 120
ggtgcagaac ctgtggaatc agccaatttg gcttgctcat ttactttaat aagggtcccat 180
aatgagtgag agtacaaagt tcaagccctg ttgagggtct gcattaaact ctcagaagta 240
tttagagtgt gccaggagcc gcgaagggtc ggttcgggtg gtggcgggaa ctgtattaga 300
gtgctaggca cggcgcgaca aagtctgtcc aacccaaaac ggtgctgagg cgttgggtgt 360
gagctccagt actcagaaaa gcattctcagc aggtactcaa cagatcctca ggggcttggg 420
ggcccagcac tggcagtgtg ggcatgaaag acataaaagg gcactacctg tgggtatttt 480
ctgttctcca aggaggaagt agcaaaaatt aggacgtctg aatatactat gttgtagcaa 540
tcccagaaca actgatgtct aacaaatacc acacaaaaca aattttttta aatttaattct 600
ta 602
```

<210> 768

<211> 671

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(671)

<223> n=A,T,C or G

<400> 768

```
tccaccgcgg tggcggccgc tctagactag tggatccact agtccagtgt ggggtgggaat 60
tcgcggcneg cgtcgacaaa aatactgcta aagtaatat tttatagatg actatttgcc 120
ttggggccag gaaaagcagc tggagtatt cacttagtac cattttttara tactaacttt 180
gcctttttcca tgcttgcttg atgcggcttg cagcactgaa gaacagtttc aattgctagc 240
caaccagaga gcatgatcaa accaaacaag ttccctgttt caggaaaaac aggttttagg 300
taactgaagg gttaccagtt actgattcca caatcttctc tgtaaaanat ttctgcctat 360
tatgcagact gggcggcttt aaanntggta aaactatnaa ataccatac aatattttta 420
nggggccccn ttatnaagct tttcaggcct tcccccttcc atagcattgg tgggatacaa 480
gaaaccttta aacagcaacn agctatcnag gcccaaaagg aaagtaattn tgatttttta 540
nagattccgn aacgaaaaaa tggttgggtt caaatacnac cttcttttta aaatggnttc 600
cttattaaac nttttttttt ttttaatttta ccccatggtc ntgatnttng ngcttccgcc 660
canaaaatng n 671
```

<210> 769

<211> 877

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(877)

<223> n=A,T,C or G

<400> 769

```

aaagctggag ctccccgcgg tggcgccgcg tctagaacta gtggatccac tagtccanng 60
ngggggaatt cgcggccgcg tcgacctcta tacctttgnt catgcagctt cctctgactg 120
ggtttgttct tcacttggct aacccctctt ttacttaagc acaccttgaa cattccctcc 180
ttccccattt ccccgagng cccctaattg acatacttct gaataacaca ggtgggtattc 240
cttccttggt ggaacctcct ggaggaagag acagatgatt aacaaatcct tccatcaacc 300
cctttgacca tgacatcaac agtgctcaa attatgggt accgtattag cctatgtcta 360
tcttgatcag aatccttacc tcggtgtatt gaaattatct atttcgtgcc tgccctcttta 420
aagtcagggg ttgccttacc tattgtctaa caccatgcag taggtaacat gcagtaggaa 480
acatggcatt aaattatctt ggttcaaata ccagttatgg tgtgtaaata cctaccaggc 540
cgtgaggcac ctgctaagca ggttgacgc atcatttgaa ttcacaccac ccttttgcaa 600
tagaacagat aggcaacaga ggctcatttg ggctaaagga tttgatggag gggaagtgcc 660
aggattccca ccaaggcctc anggccagg tccanggacc atgtctgttg tgacaactgg 720
agtgcatctc atatccctn ctctgngggg naaggtecc cncgnggaga acnnttaaaa 780
caatcatntc tngggggnnt aatgcttctt nccccagtg ggtncactgc ngccacgagt 840
cccancact agtcccangt ctgtcatgaa ccanc 877

```

<210> 770

<211> 874

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1) ... (874)

<223> n=A,T,C or G

<400> 770

```

ctggnctccc cgcggtggcg gccgctctag aactagtgga tccactagtc cagtgtggtg 60
gaattcgcg cgcgctcgac cttttcaaag gttaacttat ttaattatca cannngcaac 120
ccgatgagta ggtaacagta ttttactgat aggtaatcta aagaaggagg ctaaataaat 180
tgcccaattt cgaacagtga gaggaagaat taggattgaa acacatatag tggcttcaga 240
atctgtaacc ctcacgatgc cactactact tctttcagaa taccctttgc ctatctattc 300
tgttcctatg tcatcaaatt atacttactt taaaaagtat ttgtctttat tattttttaa 360
aaaacacagg gaagtatttc tgatcagggg cagtattggt tctgaaagac aagccagtgt 420
ttttgagggg ttctcccttg ccagtttttc tatgctgggt tattcaagtc ctaagaattg 480
tgtagctatt acagaaccgc ttttagcaaat gtgttcatt aatcaagggt atttataaca 540
aaatttcac caagtttgga gtgctctgaa aacatagcca aaatgttcgc agggctctacc 600
cctctcgtgt gtcccttttt ttttagctatt tcagaagcac actggtgcaa tatttttacga 660
aatgagtttc ttccccttac ctctgcaccc tctaagaaaa aatcattgnt gttttatgaa 720
natgaanatc ctgctatttc atatcttgat tggagctgct taattaaatg accatttttna 780
aatttgtttt gattccnngc aaaaaaagtt tnttnttgga tgtagggggc tcnnaaagnc 840
caaaaccccc caaaattttt nnttggaac ccna 874

```

<210> 771

<211> 156

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(156)

<223> n=A,T,C or G

<400> 771

```

ttaaaaaanct ggncctccccg cggtggcgccg cgctctagaa ctagtggatc cactagtcca 60
gtgtgggtgga attcgcggcc gcgtcgaccg cgagcggtcg cccttttttt tttttttttn 120
ngtttttttg aanaattcat tgggtattta ttattc 156

```

<210> 772

<211> 586

<212> DNA

<213> Homo sapiens

<220>

<221> misc_feature

<222> (1)...(586)

<223> n=A,T,C or G

<400> 772

```

ncaanctggn ctccaccgcy gtggcgcccg ctctagacta gtggatccac tagtccagt 60
tggtggaatt cgcygccgcy tcgatcacia agtgcacaca agtcnngnat ttattttatc 120
tccagatatg aaacttacc ccagctatgg tcttctatct gttatttaat ttctaggcca 180
atTTTTTcca cttgaatgtc agtattttta ttcaaagtca ccttgtccaa ataccaagtc 240
atcaacttac cctcaaatta tatcctcatt cagaaaatct acatctatta atggtagcta 300
ttttatccct gccccctgct ttttcttttt atatttaatt aatttgntca tccagcaaat 360
gcttattgag caggatttgt aggctaaaca attctanact ttaaggggac acagnttgca 420
aaacaaaatc ctgccttgna tggatactta tgnnatggng ggatacagac aatcaacata 480
atgangngca tcatatataa tggttagnan aatgataagg gnttttggga aaaaaatgca 540
ccanccaan anggattggg aagtggangg ganggtcang ggangg 586

```

<210> 773

<211> 2983

<212> DNA

<213> Homo sapiens

<400> 773

```

agagatagag tcttccctgg cattgcagga gagaatctga agggatgatg gatgcatcaa 60
aagagctgca agttctccac attgacttct tgaatcagga caacgccgtt tctcaccaca 120
catgggagtt ccaaacgagc agtcctgtgt tccggcgagg acagggtgtt cacctgcygc 180
tggtgctgaa ccagccccta caatcctacc accaactgaa actggaattc agcacagggc 240
cgaatcctag catcgccaaa cacaccctgg tgggtgctga cccgaggacg ccctcagacc 300
actacaactg gcaggcaacc cttcaaaatg agtctggcaa agaggtcaca gtggctgtca 360
ccagttcccc caatgccatc ctgggcaagt accaactaaa cgtgaaaact ggaaaccaca 420
tccttaagtc tgaagaaaac atcctatacc ttctcttcaa cccatgggtgt aaagaggaca 480
tggttttcat gcctgatgag gacgagcgca aagagtacat cctcaatgac acgggctgcc 540
attacgtggg ggctgccaga agtatcaaat gcaaaccttg gaactttggc cagtttgaga 600
aaaatgtcct ggactgctgc atttccctgc tgactgagag ctccctcaag cccacagata 660
ggagggaccc cgtgctgggtg tgcagggcca tgtgtgctat gatgagcttt gagaaaggcc 720
agggcggtgt cattgggaat tggactgggg actatgaagg tggcacagcc ccatacaagt 780
ggacaggcag tgccccgatc ctgcagcagt actacaacac gaagcaggct gtgtgctttg 840

```

```

gccagtgtg ggtgtttgct gggatcctga ctacagtgt gagagcgttg ggcattcccag 900
cacgcagtgt gacaggcttc gattcagctc acgacacaga aaggaacctc acgggtggaca 960
cctatgtgaa tgagaatggc aagaaaatca ccagtatgac ccacgactct gtctggaatt 1020
tccatgtgtg gacggatgcc tggatgaagc gaccggatct gcccaggggc tacgacggct 1080
ggcaggctgt ggacgcaacg ccgaggagc gaagccaggg tgtcttctgc tgtgggccaat 1140
caccactgac cgccatccgc aaagggtgaca tctttattgt ctatgacacc agattcgtct 1200
tctcagaagt gaatggtgac aggctcatct ggttgggtgaa gatggtgaat gggcaggagg 1260
agttacacgt aatttcaatg gagaccacaa gcacgggaa aaacatcagc accaaggcag 1320
tgggccaaga caggcggaga gatatacct atgagtacaa gtatccagaa ggctcctctg 1380
aggagaggca ggtcatggat catgccttcc tcttctcag ttctgagagg gagcacagac 1440
gacctgtaaa agagaacttt cttcacatgt cggtaacaatc agatgatgtg ctgctgggaa 1500
actctgttaa tttcaccgtg attcttaaaa ggaagaccgc tgcctacag aatgtcaaca 1560
tcttgggctc ctttgaacta cagtgtgaca ctggcaagaa gatggcaaaa ctgtgtgacc 1620
tcaataagac ctgcagatc caaggtaag tatcagaagt gactctgacc ttggactcca 1680
agacctacat caacagcctg gctatattag atgatgagcc agttatcaga ggtttcatca 1740
ttgcggaaat tgtggagtct aaggaaatca tggcctctga agtattcacg tctttccagt 1800
accctgagtt ctctatagag ttgcctaaca caggcagaat tggccagcta cttgtctgca 1860
attgtatctt caagaatacc ctggccatcc ctttgactga cgtcaagttc tctttggaaa 1920
gcctgggcat ctctcacta cagacctctg accatgggac ggtgcagcct ggtgagacca 1980
tccaatccca aataaaatgc acccaataa aaactggacc caagaaattt atcgtcaagt 2040
taagttccaa acaagtgaat gagattaatg ctcagaagat tgttctcatc accaagtagc 2100
cttgtctgat gctgtggagc cttagttgag atttcagcat tctctacctt gtgcttagct 2160
ttcagattat ggtgatttaa atttgatgac ttatatgagg gcagattcaa gagccagcag 2220
gtcaaaaagg ccaacacaac cataagcagc cagaccacac aggccaggtc ctgtgctatc 2280
acagggtcac ctcttttaca gttagaaaca ccagccgagg ccacagaatc ccatcccttt 2340
cctgagtcac ggccctcaaaa atcagggcca ccattgtctc aattcaaata catagatttc 2400
gaagccacag agtctctccc tggagcagca gactatgggc agcccagtgc tgccacctgc 2460
tgacgacctc tgagaagctg ccatactctc aggccatggg ttcaccagcc ctgaaggcac 2520
ctgtcaactg gagtgtctct tcagcactgg gatgggcctg atagaagtgc attctcctcc 2580
tattgctctc attctcctct ctctatccct gaaatccagg aagtcctctc cctggtgctc 2640
caagcagttt gaagcccaat ctgcaaggac atttctcaag ggccatgtgg ttttgagac 2700
aaccctgtcc tcaggcctga actcaccata gagaccatg tcagcaaacg gtgaccagca 2760
aatcctcttc ccttattcta aagctgcccc ttgggagact ccagggagaa ggcattgctt 2820
cctcctgggt gtgaactctt tctttgggat tccatccact atcctggcaa ctcaaggctg 2880
cttctgttaa ctgaagcctg ctcttctctg ttctgccctc cagagatttg ctcaaatgat 2940
caataagctt taaattaaac tctacttcaa gaaaaaaaaa ccg 2983

```

<210> 774

<211> 3064

<212> DNA

<213> Homo sapiens

<400> 774

```

aattctaaaa atgtttttgc aagcttgcac gcctgcaggt gcagcggccg ccagtgtgat 60
ggatatctgc agaattcggc ttgcgctcag ctggaattcc gcagagatag agtcttccct 120
ggcattgcag gagagaatct gaagggatga tggatgcac aaaagagctg caagttctcc 180
acattgactt cttgaatcag gacaacgcc tttctacca cacatgggag ttccaaacga 240
gcagtcctgt gttccggcga ggacagggtg ttcacctgcg gctggtgctg aaccagcccc 300
tacaatccta ccaccaactg aaactggaat tcagcacagg gccgaatcct agcatcgcca 360
aacacaccct ggtggtgctc gacccgagga cgccctcaga ccactacaac tggcaggcaa 420
cccttcaaaa tgagtctggc aaagaggtea cagtggctgt caccagttcc cccaatgcc 480

```

tccctgggcaa gtaccaacta aacgtgaaaa ctggaaacca catccttaag tctgaagaaa 540
 acatcctata ccttctcttc aacccatggt gtaaagagga catggttttc atgcctgatg 600
 aggacgagcg caaagagtac atcctcaatg acacgggctg ccattacgtg ggggctgcca 660
 gaagtatcaa atgcaaacc tggaaactttg gtcagtttga gaaaaatgtc ctggactgct 720
 gcatttccct gctgactgag agctccctca agcccacaga taggagggac cccgtgctgg 780
 tgtgcagggc catgtgtgct atgatgagct ttgagaaagg ccagggcgctg ctcatgtgga 840
 attggactgg ggactacgaa ggtggcacag cccatacaa gtggacaggc agtgccccga 900
 tcctgcagca gtactacaac acgaagcagg ctgtgtgctt tggccagtgc tgggtgtttg 960
 ctgggatcct gactacagtg ctgagagcgt tgggcatccc agcacgcagt gtgacaggct 1020
 tcgattcagc tcacgacaca gaaaggaacc tcacggtgga cacctatgtg aatgagaatg 1080
 gcgagaaaat caccagtatg acccacgact ctgtctggaa tttccatgtg tggacggatg 1140
 cctggatgaa gcgaccctac gacggctggc aggtctgtgga cgcaacgcgc caggagcgaa 1200
 gccaggggtg cttctgctgt gggccatcac cactgaccgc catccgcaa ggtgacatct 1260
 ttattgtcta tgacaccaga ttctcttctt cagaagtga tgggtgacagg ctcatctggt 1320
 tgggtgaagat ggtgaatggg caggaggagt tacacgtaat ttcaatggag accacaagca 1380
 tcgggaaaaa catcagcacc aaggcagtgg gccaaagacag gcggagagat atcacctatg 1440
 agtacaagta tccagaaggc tcctctgagg agaggcagg catggatcat gccttctctc 1500
 ttctcagttc tgagagggag cacagacagc ctgtaaaaga gaactttctt cacatgtcgg 1560
 tacaatcaga tgatgtgctg ctgggaaact ctgttaattt caccgtgatt cttaaaagga 1620
 agaccgctgc cctacagaat gtcaacatct tgggctcctt tgaactacag ttgtacactg 1680
 gcaagaagat ggcaaaactg tgtgacctca ataagacctc gcagatccaa ggtcaagtat 1740
 cagaagtgc tctgaccttg gactccaaga cctacatcaa cagcctggct atattagatg 1800
 atgagccagt tatcagaggt tcatcattg cggaaattgt ggagtctaag gaaatcatgg 1860
 cctctgaagt attcacgtca aaccagtacc ctgagttctc tatagagttg cctaacacag 1920
 gcagaattgg ccagctactt gtctgcaatt gtatcttcaa gaataccctg gccatccctt 1980
 tgactgacgt caagtctctt ttggaaagcc tgggcatctc ctactacag acctctgacc 2040
 atgggacggt gcagcctggt gagaccatcc aatcccaaat aaaatgcacc ccaataaaaa 2100
 ctggacccaa gaaatttata gtcaagttaa gttccaaaca agtgaaagag attaatgctc 2160
 agaagattgt tctcatcacc aagtagcctt gtctgatgct gtggagcctt agttgagatt 2220
 tcagcatttc ctacettgtg cttagctttc agattatgga tgattaaatt tgatgactta 2280
 tatgagggca gattcaagag ccagcaggtc aaaaaggcca acacaacat aagcagccag 2340
 accacaagg ccaggtcctg tgctatcaca ggggtcacctc ttttacagtt agaaacacca 2400
 gccgaggcca cagaatccca tccctttcct gagtcatggc ctcaaaaatc agggccacca 2460
 ttgtctcaat tcaaatccat agatttcgaa gccacagagc tcttccctgg agcagcagac 2520
 tatgggcagc ccagtgtgct cacctgctga cgacccttga gaagctgcca tatcttcagg 2580
 ccatgggttc accagccctg aaggcacctg tcaactggag tgctctctca gcactgggat 2640
 gggcctgata gaagtgcatt ctccctcctat tgccctcatt ctccctctctc tatccctgaa 2700
 atccaggaag tccctctcct ggtgctccaa gcagtttgaa gcccaatctg caaggacatt 2760
 tctcaagggc catgtggttt tgcagacaac cctgtcctca ggctgaact caccatagag 2820
 acccatgtca gcaaacgggt accagcaaat cctcttccct tattctaaag ctgccccttg 2880
 ggagactcca gggagaaggc attgcttccct ccctgggtgtg aactctttct ttggtattcc 2940
 atccactatc ctggcaactc aaggctgctt ctgttaactg aagcctgctc cttctgttct 3000
 tgccctccag agatttgctc aaatgatcaa taagctttaa attaaaccgg aatccgcgga 3060
 attc 3064

<210> 775

<211> 684

<212> PRT

<213> Homo sapiens

<400> 775

Ser Lys Thr Tyr Ile Asn Ser Leu Ala Ile Leu Asp Asp Glu Pro Val
 545 550 555 560
 Ile Arg Gly Phe Ile Ile Ala Glu Ile Val Glu Ser Lys Glu Ile Met
 565 570 575

Ala Ser Glu Val Phe Thr Ser Phe Gln Tyr Pro Glu Phe Ser Ile Glu
 580 585 590

Leu Pro Asn Thr Gly Arg Ile Gly Gln Leu Leu Val Cys Asn Cys Ile
 595 600 605

Phe Lys Asn Thr Leu Ala Ile Pro Leu Thr Asp Val Lys Phe Ser Leu
 610 615 620

Glu Ser Leu Gly Ile Ser Ser Leu Gln Thr Ser Asp His Gly Thr Val
 625 630 635 640

Gln Pro Gly Glu Thr Ile Gln Ser Gln Ile Lys Cys Thr Pro Ile Lys
 645 650 655

Thr Gly Pro Lys Lys Phe Ile Val Lys Leu Ser Ser Lys Gln Val Lys
 660 665 670

Glu Ile Asn Ala Gln Lys Ile Val Leu Ile Thr Lys
 675 680

<210> 776

<211> 679

<212> PRT

<213> Homo sapiens

<400> 776

Met Met Asp Ala Ser Lys Glu Leu Gln Val Leu His Ile Asp Phe Leu
 5 10 15

Asn Gln Asp Asn Ala Val Ser His His Thr Trp Glu Phe Gln Thr Ser
 20 25 30

Ser Pro Val Phe Arg Arg Gly Gln Val Phe His Leu Arg Leu Val Leu
 35 40 45

Asn Gln Pro Leu Gln Ser Tyr His Gln Leu Lys Leu Glu Phe Ser Thr
 50 55 60

Gly Pro Asn Pro Ser Ile Ala Lys His Thr Leu Val Val Leu Asp Pro
 65 70 75 80

Arg Thr Pro Ser Asp His Tyr Asn Trp Gln Ala Thr Leu Gln Asn Glu
 85 90 95

Gly Asp Ile Phe Ile Val Tyr Asp Thr Arg Phe Val Phe Ser Glu Val
 370 375 380
 Asn Gly Asp Arg Leu Ile Trp Leu Val Lys Met Val Asn Gly Gln Glu
 385 390 395 400
 Glu Leu His Val Ile Ser Met Glu Thr Thr Ser Ile Gly Lys Asn Ile
 405 410 415
 Ser Thr Lys Ala Val Gly Gln Asp Arg Arg Arg Asp Ile Thr Tyr Glu
 420 425 430
 Tyr Lys Tyr Pro Glu Gly Ser Ser Glu Glu Arg Gln Val Met Asp His
 435 440 445
 Ala Phe Leu Leu Leu Ser Ser Glu Arg Glu His Arg Gln Pro Val Lys
 450 455 460
 Glu Asn Phe Leu His Met Ser Val Gln Ser Asp Asp Val Leu Leu Gly
 465 470 475 480
 Asn Ser Val Asn Phe Thr Val Ile Leu Lys Arg Lys Thr Ala Ala Leu
 485 490 495
 Gln Asn Val Asn Ile Leu Gly Ser Phe Glu Leu Gln Leu Tyr Thr Gly
 500 505 510
 Lys Lys Met Ala Lys Leu Cys Asp Leu Asn Lys Thr Ser Gln Ile Gln
 515 520 525
 Gly Gln Val Ser Glu Val Thr Leu Thr Leu Asp Ser Lys Thr Tyr Ile
 530 535 540
 Asn Ser Leu Ala Ile Leu Asp Asp Glu Pro Val Ile Arg Gly Phe Ile
 545 550 555 560
 Ile Ala Glu Ile Val Glu Ser Lys Glu Ile Met Ala Ser Glu Val Phe
 565 570 575
 Thr Ser Asn Gln Tyr Pro Glu Phe Ser Ile Glu Leu Pro Asn Thr Gly
 580 585 590
 Arg Ile Gly Gln Leu Leu Val Cys Asn Cys Ile Phe Lys Asn Thr Leu
 595 600 605
 Ala Ile Pro Leu Thr Asp Val Lys Phe Ser Leu Glu Ser Leu Gly Ile
 610 615 620
 Ser Ser Leu Gln Thr Ser Asp His Gly Thr Val Gln Pro Gly Glu Thr
 625 630 635 640


```

cagctttgaa tacactataa actcagtggc tgaaggagga aatttttagaa ggaagctact 5160
aaaagatcta atttgaaaaa ctacaaaagc attaactaaa aaagtttatt ttccttttgt 5220
ctgggcagta gtgaaaataa ctactcacia cattcactat gtttgcaagg aattaacaca 5280
aataaaagat gccttttttac ttaaacgcca agacagaaaaa cttgccaat actgagaagc 5340
aacttgcatt agagagggaa ctgttaaatg ttttcaaccc agttcatctg gtggatgttt 5400
ttgcaggtta ctctgagaat tttgcttatg aaaaatcatt attttttagtg tagttcacaa 5460
taatgtattg aacatacttc taatcaaagg tgctatgtcc ttgtgtatgg tactaaatgt 5520
gtcctgtgta cttttgcaca actgagaatc ctgcggcttg gtttaatgag tgtgttcatg 5580
aaataaataa tggaggaatt gtcaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 5640
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa                                     5668

```

<210> 778

<211> 1095

<212> PRT

<213> Homo sapiens

<400> 778

```

Met Arg Asn Arg Arg Asn Asp Thr Leu Asp Ser Thr Arg Thr Leu Tyr
      5                      10                      15

Ser Ser Ala Ser Arg Ser Thr Asp Leu Ser Tyr Ser Glu Ser Asp Leu
      20                      25                      30

Val Asn Phe Ile Gln Ala Asn Phe Lys Lys Arg Glu Cys Val Phe Phe
      35                      40                      45

Thr Lys Asp Ser Lys Ala Thr Glu Asn Val Cys Lys Cys Gly Tyr Ala
      50                      55                      60

Gln Ser Gln His Met Glu Gly Thr Gln Ile Asn Gln Ser Glu Lys Trp
      65                      70                      75                      80

Asn Tyr Lys Lys His Thr Lys Glu Phe Pro Thr Asp Ala Phe Gly Asp
      85                      90                      95

Ile Gln Phe Glu Thr Leu Gly Lys Lys Gly Lys Tyr Ile Arg Leu Ser
      100                     105                     110

Cys Asp Thr Asp Ala Glu Ile Leu Tyr Glu Leu Leu Thr Gln His Trp
      115                     120                     125

His Leu Lys Thr Pro Asn Leu Val Ile Ser Val Thr Gly Gly Ala Lys
      130                     135                     140

Asn Phe Ala Leu Lys Pro Arg Met Arg Lys Ile Phe Ser Arg Leu Ile
      145                     150                     155                     160

Tyr Ile Ala Gln Ser Lys Gly Ala Trp Ile Leu Thr Gly Gly Thr His
      165                     170                     175

Tyr Gly Leu Thr Lys Tyr Ile Gly Glu Val Val Arg Asp Asn Thr Ile

```

				180						185						190
Ser	Arg	Ser	Ser	Glu	Glu	Asn	Ile	Val	Ala	Ile	Gly	Ile	Ala	Ala	Trp	
		195					200					205				
Gly	Met	Val	Ser	Asn	Arg	Asp	Thr	Leu	Ile	Arg	Asn	Cys	Asp	Ala	Glu	
	210					215					220					
Gly	Tyr	Phe	Leu	Ala	Gln	Tyr	Leu	Met	Asp	Asp	Phe	Thr	Arg	Asp	Pro	
225					230				235						240	
Leu	Tyr	Ile	Leu	Asp	Asn	Asn	His	Thr	His	Leu	Leu	Leu	Val	Asp	Asn	
				245					250					255		
Gly	Cys	His	Gly	His	Pro	Thr	Val	Glu	Ala	Lys	Leu	Arg	Asn	Gln	Leu	
			260					265					270			
Glu	Lys	His	Ile	Ser	Glu	Arg	Thr	Ile	Gln	Asp	Ser	Asn	Tyr	Gly	Gly	
		275					280					285				
Lys	Ile	Pro	Ile	Val	Cys	Phe	Ala	Gln	Gly	Gly	Gly	Lys	Glu	Thr	Leu	
	290					295					300					
Lys	Ala	Ile	Asn	Thr	Ser	Ile	Lys	Asn	Lys	Ile	Pro	Cys	Val	Val	Val	
305					310					315					320	
Glu	Gly	Ser	Gly	Arg	Ile	Ala	Asp	Val	Ile	Ala	Ser	Leu	Val	Glu	Val	
				325					330					335		
Glu	Asp	Ala	Pro	Thr	Ser	Ser	Ala	Val	Lys	Glu	Lys	Leu	Val	Arg	Phe	
			340					345					350			
Leu	Pro	Arg	Thr	Val	Ser	Arg	Leu	Ser	Glu	Glu	Glu	Thr	Glu	Ser	Trp	
		355					360					365				
Ile	Lys	Trp	Leu	Lys	Glu	Ile	Leu	Glu	Cys	Ser	His	Leu	Leu	Thr	Val	
	370					375					380					
Ile	Lys	Met	Glu	Glu	Ala	Gly	Asp	Glu	Ile	Val	Ser	Asn	Ala	Ile	Ser	
385					390					395					400	
Tyr	Ala	Leu	Tyr	Lys	Ala	Phe	Ser	Thr	Ser	Glu	Gln	Asp	Lys	Asp	Asn	
				405					410					415		
Trp	Asn	Gly	Gln	Leu	Lys	Leu	Leu	Leu	Glu	Trp	Asn	Gln	Leu	Asp	Leu	
			420					425					430			
Ala	Asn	Asp	Glu	Ile	Phe	Thr	Asn	Asp	Arg	Arg	Trp	Glu	Ser	Ala	Asp	
	435						440					445				
Leu	Gln	Glu	Val	Met	Phe	Thr	Ala	Leu	Ile	Lys	Asp	Arg	Pro	Lys	Phe	

725										730					735				
Ile	Ala	Phe	Leu	Leu	Leu	Phe	Ala	Tyr	Val	Leu	Leu	Met	Asp	Phe	His				
			740					745					750						
Ser	Val	Pro	His	Pro	Pro	Glu	Leu	Val	Leu	Tyr	Ser	Leu	Val	Phe	Val				
		755					760					765							
Leu	Phe	Cys	Asp	Glu	Val	Arg	Gln	Trp	Tyr	Val	Asn	Gly	Val	Asn	Tyr				
	770					775					780								
Phe	Thr	Asp	Leu	Trp	Asn	Val	Met	Asp	Thr	Leu	Gly	Leu	Phe	Tyr	Phe				
785					790					795					800				
Ile	Ala	Gly	Ile	Val	Phe	Arg	Leu	His	Ser	Ser	Asn	Lys	Ser	Ser	Leu				
			805						810					815					
Tyr	Ser	Gly	Arg	Val	Ile	Phe	Cys	Leu	Asp	Tyr	Ile	Ile	Phe	Thr	Leu				
		820					825						830						
Arg	Leu	Ile	His	Ile	Phe	Thr	Val	Ser	Arg	Asn	Leu	Gly	Pro	Lys	Ile				
	835						840					845							
Ile	Met	Leu	Gln	Arg	Met	Leu	Ile	Asp	Val	Phe	Phe	Phe	Leu	Phe	Leu				
	850					855					860								
Phe	Ala	Val	Trp	Met	Val	Ala	Phe	Gly	Val	Ala	Arg	Gln	Gly	Ile	Leu				
865					870					875					880				
Arg	Gln	Asn	Glu	Gln	Arg	Trp	Arg	Trp	Ile	Phe	Arg	Ser	Val	Ile	Tyr				
			885						890					895					
Glu	Pro	Tyr	Leu	Ala	Met	Phe	Gly	Gln	Val	Pro	Ser	Asp	Val	Asp	Gly				
		900						905					910						
Thr	Thr	Tyr	Asp	Phe	Ala	His	Cys	Thr	Phe	Thr	Gly	Asn	Glu	Ser	Lys				
		915					920					925							
Pro	Leu	Cys	Val	Glu	Leu	Asp	Glu	His	Asn	Leu	Pro	Arg	Phe	Pro	Glu				
	930					935					940								
Trp	Ile	Thr	Ile	Pro	Leu	Val	Cys	Ile	Tyr	Met	Leu	Ser	Thr	Asn	Ile				
945				950					955						960				
Leu	Leu	Val	Asn	Leu	Leu	Val	Ala	Met	Phe	Gly	Tyr	Thr	Val	Gly	Thr				
			965						970					975					
Val	Gln	Glu	Asn	Asn	Asp	Gln	Val	Trp	Lys	Phe	Gln	Arg	Tyr	Phe	Leu				
			980					985					990						
Val	Gln	Glu	Tyr	Cys	Ser	Arg	Leu	Asn	Ile	Pro	Phe	Pro	Phe	Ile	Val				


```

cagctgaagc ttctgctgga gtggaaccag ctggacttag ccaatgatga gattttcacc 1440
aatgaccgcc gatgggagtc tgctgacctt caagaagtca tgtttacggc tctcataaag 1500
gacagaccca agtttgccg cctctttctg gagaatggct tgaacctacg gaagtttctc 1560
acccatgatg tcctcactga actcttctcc aaccacttca gcacgcttgt gtaccggaat 1620
ctgcagatcg ccaagaattc ctataatgat gccctcctca cgtttgtctg gaaactgggt 1680
gcgaacttcc gaagaggctt ccggaaggaa gacagaaatg gccgggacga gatggacata 1740
gaactccacg acgtgtctcc tattactcgg caccctctgc aagctctctt catctgggcc 1800
attcttcaga ataagaagga actctccaaa gtcatctggg agcagaccag gggctgcact 1860
ctggcagccc tgggagccag caagcttctg aagactctgg ccaaagtga gaacgacatc 1920
aatgctgctg gggagtccga ggagctggct aatgagtaag agaccggggc tgttgagctg 1980
ttcactgagt gttacagcag cgatgaagac ttggcagAAC agctgctggg ctattcctgt 2040
gaagcttggg gtggaagcaa ctgtctggag ctggcgggtg aggccacaga ccagcatttc 2100
atcgcccagc ctgggggtcca gaattttctt tctaagcaat ggtatggaga gatttcccg 2160
gacaccaaga actggaagat tatcctgtgt ctgtttatta tacccttggg gggctgtggc 2220
tttgtatcat ttaggaagaa acctgtcgac aagcacaaga agctgctttg gtactatgtg 2280
gcgttcttca cctccccctt cgtgggtctt tcttggaaatg tgggtcttcta catcgcttc 2340
ctcctgctgt ttgctacgt gctgtctcat gattttccatt cgggtgccaca ccccccgag 2400
ctggctcctg actcgctggg ctttgtctct ttctgtgatg aagtgagaca gtggtacgta 2460
aatgggggtg attattttac tgacctgtgg aatgtgatgg acacgctggg gcttttttac 2520
ttcatagcag gaattgtatt tcggctccac tcttctaata aaagctcttt gtattctgga 2580
cgagtcattt tctgtctgga ctacattatt ttcactctaa gattgatcca catttttact 2640
gtaagcagaa acttaggacc caagattata atgctgcaga ggatgctgat cgatgtgttc 2700
ttcttctgt tctcttttg gcgwtggatg gtggcctttg gcgtggccag gcaagggatc 2760
cttaggcaga atgagcagcg ctggaggtgg atattccgtt cggtcactta cgagccctac 2820
ctggccatgt tcggccaggt gccagtgac gtggatggta ccacgtatga ctttgccac 2880
tgcacettca ctgggaatga gtccaagcca ctgtgtgtgg agctggatga gcacaacctg 2940
ccccggttcc ccgagtggat caccatcccc ctgggtgtgca tctacatgtt atccaccaac 3000
atcctgctgg tcaacctgct ggtcgccatg tttggctaca cgggtgggcac cgtccaggag 3060
aacaatgacc aggtctggaa gttccagagg tacttcttgg tgcaggagta ctgcagccgc 3120
ctcaatatcc ccttccccct catcgtcttc gcttacttct acatgggtgg gaagaagtgc 3180
ttcaagtgtt gctgcaagga gaaaaacatg gagtcttctg tctgctgttt caaaaatgaa 3240
gacaatgaga ctctggcatg ggagggtgtc atgaagggaa actaccttgt caagatcaac 3300
acaaaagcca acgacacctc agaggaaatg aggcacgat ttagacaact ggatacaaag 3360
cttaatgatc tcaagggtct tctgaaagag attgctaata aaatcaaata aaactgtatg 3420
aactctaata gagaaaaatc taattatagc aagatcatat taaggaatgc tgatgaacaa 3480
ttttgctatc gactactaaa tgagagattt tcagaccctt gggtacatgg tggatgattt 3540
taaatcacc tagtgtgctg agaccttgag aataaagtgt gaagggcgaa ttctgcagat 3600
atccatcaca ctggcgggccg ctcgagcatg catctagag 3639

```

<210> 780

<211> 1095

<212> PRT

<213> Homo sapiens

<220>

<221> VARIANT

<222> (1)...(1095)

<223> Xaa = Any Amino Acid

<400> 780

Met Arg Asn Arg Arg Asn Asp Thr Leu Asp Ser Thr Arg Thr Leu Tyr

5					10					15					
Ser	Ser	Ala	Ser	Arg	Ser	Thr	Asp	Leu	Ser	Tyr	Ser	Glu	Ser	Asp	Leu
			20					25					30		
Val	Asn	Phe	Ile	Gln	Ala	Asn	Phe	Lys	Lys	Arg	Glu	Cys	Val	Phe	Phe
		35					40					45			
Thr	Lys	Asp	Ser	Lys	Ala	Thr	Glu	Asn	Val	Cys	Lys	Cys	Gly	Tyr	Ala
	50					55					60				
Gln	Ser	Gln	His	Met	Glu	Gly	Thr	Gln	Ile	Asn	Gln	Ser	Glu	Lys	Trp
65					70					75					80
Asn	Tyr	Lys	Lys	His	Thr	Lys	Glu	Phe	Pro	Thr	Asp	Ala	Phe	Gly	Asp
				85					90					95	
Ile	Gln	Phe	Glu	Thr	Leu	Gly	Lys	Lys	Gly	Lys	Tyr	Ile	Arg	Leu	Ser
			100					105					110		
Cys	Asp	Thr	Asp	Ala	Glu	Ile	Leu	Tyr	Glu	Leu	Leu	Thr	Gln	His	Trp
	115						120					125			
His	Leu	Lys	Thr	Pro	Asn	Leu	Val	Ile	Ser	Val	Thr	Gly	Gly	Ala	Lys
	130					135					140				
Asn	Phe	Ala	Leu	Lys	Pro	Arg	Met	Arg	Lys	Ile	Phe	Ser	Arg	Leu	Ile
145					150					155					160
Tyr	Ile	Ala	Gln	Ser	Lys	Gly	Ala	Trp	Ile	Leu	Thr	Gly	Gly	Thr	His
			165					170						175	
Tyr	Gly	Leu	Met	Lys	Tyr	Ile	Gly	Glu	Val	Val	Arg	Asp	Asn	Thr	Ile
		180						185					190		
Ser	Arg	Ser	Ser	Glu	Glu	Asn	Ile	Val	Ala	Ile	Gly	Ile	Ala	Ala	Trp
		195					200					205			
Gly	Met	Val	Ser	Asn	Arg	Asp	Thr	Leu	Ile	Arg	Asn	Cys	Asp	Ala	Glu
	210					215					220				
Gly	Tyr	Phe	Leu	Ala	Gln	Tyr	Leu	Met	Asp	Asp	Phe	Thr	Arg	Asp	Pro
225				230						235					240
Leu	Tyr	Ile	Leu	Asp	Asn	Asn	His	Thr	His	Leu	Leu	Leu	Val	Asp	Asn
			245						250					255	
Gly	Cys	His	Gly	His	Pro	Thr	Val	Glu	Ala	Lys	Leu	Arg	Asn	Gln	Leu
		260						265					270		
Glu	Lys	Tyr	Ile	Ser	Glu	Arg	Thr	Ile	Gln	Asp	Ser	Asn	Tyr	Gly	Gly

545		550		555		560
Leu Gln Asn Lys Lys Glu Leu Ser Lys Val Ile Trp Glu Gln Thr Arg						
		565		570		575
Gly Cys Thr Leu Ala Ala Leu Gly Ala Ser Lys Leu Leu Lys Thr Leu						
		580		585		590
Ala Lys Val Lys Asn Asp Ile Asn Ala Ala Gly Glu Ser Glu Glu Leu						
		595		600		605
Ala Asn Glu Tyr Glu Thr Arg Ala Val Glu Leu Phe Thr Glu Cys Tyr						
		610		615		620
Ser Ser Asp Glu Asp Leu Ala Glu Gln Leu Leu Val Tyr Ser Cys Glu						
		625		630		635
				635		640
Ala Trp Gly Gly Ser Asn Cys Leu Glu Leu Ala Val Glu Ala Thr Asp						
				645		650
						655
Gln His Phe Ile Ala Gln Pro Gly Val Gln Asn Phe Leu Ser Lys Gln						
		660		665		670
Trp Tyr Gly Glu Ile Ser Arg Asp Thr Lys Asn Trp Lys Ile Ile Leu						
		675		680		685
Cys Leu Phe Ile Ile Pro Leu Val Gly Cys Gly Phe Val Ser Phe Arg						
		690		695		700
Lys Lys Pro Val Asp Lys His Lys Lys Leu Leu Trp Tyr Tyr Val Ala						
		705		710		715
				715		720
Phe Phe Thr Ser Pro Phe Val Val Phe Ser Trp Asn Val Val Phe Tyr						
				725		730
						735
Ile Ala Phe Leu Leu Leu Phe Ala Tyr Val Leu Leu Met Asp Phe His						
		740		745		750
Ser Val Pro His Pro Pro Glu Leu Val Leu Tyr Ser Leu Val Phe Val						
		755		760		765
Leu Phe Cys Asp Glu Val Arg Gln Trp Tyr Val Asn Gly Val Asn Tyr						
		770		775		780
Phe Thr Asp Leu Trp Asn Val Met Asp Thr Leu Gly Leu Phe Tyr Phe						
		785		790		795
				795		800
Ile Ala Gly Ile Val Phe Arg Leu His Ser Ser Asn Lys Ser Ser Leu						
		805		810		815
Tyr Ser Gly Arg Val Ile Phe Cys Leu Asp Tyr Ile Ile Phe Thr Leu						

1090

1095

<210> 781

<211> 15

<212> PRT

<213> Homo sapiens

<400> 781

Arg	Met	Pro	Thr	Val	Leu	Gln	Cys	Val	Asn	Val	Ser	Val	Val	Ser
					5				10				15	

<210> 782

<211> 45

<212> DNA

<213> Homo sapiens

<400> 782

agaatgccta	ccgtgctgca	gtgCGTgaac	gtgtCGgtgg	tgtct	45
------------	------------	------------	------------	-------	----

<210> 783

<211> 45

<212> DNA

<213> Homo sapiens

<400> 783

gagccaggga	gccagatggT	ggaggccagc	ctctccgtac	ggcac	45
------------	------------	------------	------------	-------	----

<210> 784

<211> 45

<212> DNA

<213> Homo sapiens

<400> 784

gaggccgacc	aagagccagg	gagccagatg	gtggaggcca	gcctc	45
------------	------------	------------	------------	-------	----

<210> 785

<211> 45

<212> DNA

<213> Homo sapiens

<400> 785

ggcctgcaca	gtcttgaggc	cgaccaagag	ccaggggagcc	agatg	45
------------	------------	------------	-------------	-------	----

<210> 786

<211> 45

<212> DNA

<213> Homo sapiens

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

<210> 793
 <211> 51
 <212> DNA
 <213> Homo sapiens

<400> 793
 tcgcagccct ggcaggcggc actgggtcatg gaaaacgaat tgttctgctc g 51

<210> 794
 <211> 45
 <212> DNA
 <213> Homo sapiens

<400> 794
 atcagcattg cttcgcagtg ccctaccgcg gggaaactctt gcctc 45

<210> 795
 <211> 45
 <212> DNA
 <213> Homo sapiens

<400> 795
 tccgtgtccg agtctgacac catccggagc atcagcattg cttcg 45

<210> 796
 <211> 45
 <212> DNA
 <213> Homo sapiens

<400> 796
 atcaagttgg acgaatccgt gtccgagtct gacaccatcc ggagc 45

<210> 797
 <211> 45
 <212> DNA
 <213> Homo sapiens

<400> 797
 aacgacctca tgctcatcaa gttggacgaa tccgtgtccg agtct 45

<210> 798
 <211> 45
 <212> DNA
 <213> Homo sapiens

<400> 798
 agacccttgc tcgctaacga cctcatgctc atcaagttgg acgaa 45

<210> 799
 <211> 15
 <212> PRT

<213> Homo sapiens

<400> 799

Glu Pro Gly Ser Gln Met Val Glu Ala Ser Leu Ser Val Arg His
 5 10 15

<210> 800

<211> 15

<212> PRT

<213> Homo sapiens

<400> 800

Glu Ala Asp Gln Glu Pro Gly Ser Gln Met Val Glu Ala Ser Leu
 5 10 15

<210> 801

<211> 15

<212> PRT

<213> Homo sapiens

<400> 801

Gly Leu His Ser Leu Glu Ala Asp Gln Glu Pro Gly Ser Gln Met
 5 10 15

<210> 802

<211> 15

<212> PRT

<213> Homo sapiens

<400> 802

Tyr Thr Ile Gly Leu Gly Leu His Ser Leu Glu Ala Asp Gln Glu
 5 10 15

<210> 803

<211> 14

<212> PRT

<213> Homo sapiens

<400> 803

Phe Gln Asn Ser Tyr Thr Ile Gly Leu Gly Leu His Ser Leu
 5 10

<210> 804

<211> 15

<212> PRT

<213> Homo sapiens

<400> 804

Leu Ser Ala Ala His Cys Phe Gln Asn Ser Tyr Thr Ile Gly Leu
 5 10 15

<210> 805

<211> 15

<212> PRT

<213> Homo sapiens

<400> 805

His Pro Gln Trp Val Leu Ser Ala Ala His Cys Phe Gln Asn Ser
 5 10 15

<210> 806

<211> 15

<212> PRT

<213> Homo sapiens

<400> 806

Ser Gly Val Leu Val His Pro Gln Trp Val Leu Ser Ala Ala His
 5 10 15

<210> 807

<211> 15

<212> PRT

<213> Homo sapiens

<400> 807

Asn Glu Leu Phe Cys Ser Gly Val Leu Val His Pro Gln Trp Val
 5 10 15

<210> 808

<211> 15

<212> PRT

<213> Homo sapiens

<400> 808

Ala Leu Val Met Glu Asn Glu Leu Phe Cys Ser Gly Val Leu Val
 5 10 15

<210> 809

<211> 17

<212> PRT

<213> Homo sapiens

<400> 809

Ser Gln Pro Trp Gln Ala Ala Leu Val Met Glu Asn Glu Leu Phe Cys
 5 10 15

Ser

<210> 810

<211> 15

<212> PRT

<213> Homo sapiens

<400> 810

Ile Ser Ile Ala Ser Gln Cys Pro Thr Ala Gly Asn Ser Cys Leu
 5 10 15

<210> 811

<211> 15

<212> PRT

<213> Homo sapiens

<400> 811

Ser Val Ser Glu Ser Asp Thr Ile Arg Ser Ile Ser Ile Ala Ser
 5 10 15

<210> 812

<211> 15

<212> PRT

<213> Homo sapiens

<400> 812

Ile Lys Leu Asp Glu Ser Val Ser Glu Ser Asp Thr Ile Arg Ser
 5 10 15

<210> 813

<211> 15

<212> PRT

<213> Homo sapiens

<400> 813

Asn Asp Leu Met Leu Ile Lys Leu Asp Glu Ser Val Ser Glu Ser
 5 10 15

<210> 814

<211> 15

<212> PRT

Trp Asn Gly Gln Leu Lys Leu Leu Leu Glu Trp Asn Gln Leu Asp Leu
 420 425 430
 Ala Asn Asp Glu Ile Phe Thr Asn Asp Arg Arg Trp Glu Ser Ala Asp
 435 440 445
 Leu Gln Glu Val Met Phe Thr Ala Leu Ile Lys Asp Arg Pro Lys Phe
 450 455 460
 Val Arg Leu Phe Leu Glu Asn Gly Leu Asn Leu Arg Lys Phe Leu Thr
 465 470 475 480
 His Asp Val Leu Thr Glu Leu Phe Ser Asn His Phe Ser Thr Leu Val
 485 490 495
 Tyr Arg Asn Leu Gln Ile Ala Lys Asn Ser Tyr Asn Asp Ala Leu Leu
 500 505 510
 Thr Phe Val Trp Lys Leu Val Ala Asn Phe Arg Arg Gly Phe Arg Lys
 515 520 525
 Glu Asp Arg Asn Gly Arg Asp Glu Met Asp Ile Glu Leu His Asp Val
 530 535 540
 Ser Pro Ile Thr Arg His Pro Leu Gln Ala Leu Phe Ile Trp Ala Ile
 545 550 555 560
 Leu Gln Asn Lys Lys Glu Leu Ser Lys Val Ile Trp Glu Gln Thr Arg
 565 570 575
 Gly Cys Thr Leu Ala Ala Leu Gly Ala Ser Lys Leu Leu Lys Thr Leu
 580 585 590
 Ala Lys Val Lys Asn Asp Ile Asn Ala Ala Gly Glu Ser Glu Glu Leu
 595 600 605
 Ala Asn Glu Tyr Glu Thr Arg Ala Val Glu Leu Phe Thr Glu Cys Tyr
 610 615 620
 Ser Ser Asp Glu Asp Leu Ala Glu Gln Leu Leu Val Tyr Ser Cys Glu
 625 630 635 640
 Ala Trp Gly Gly Leu Glu His His His His His His
 645 650

<210> 819

<211> 132

<212> PRT

<213> Homo sapien

<400> 819

```

Thr Ala Ala Ser Asp Asn Phe Gln Leu Ser Gln Gly Gly Gln Gly Phe
1      5      10      15
Ala Ile Pro Ile Gly Gln Ala Met Ala Ile Ala Gly Gln Ile Arg Ser
20     25     30
Gly Gly Gly Ser Pro Thr Val His Ile Gly Pro Thr Ala Phe Leu Gly
35     40     45
Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val Gln Arg Val
50     55     60
Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr Gly Asp Val
65     70     75     80
Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr Ala Met Ala
85     90     95
Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser Val Asn Trp
100    105    110
Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr Leu Ala Glu
115    120    125
Gly Pro Pro Ala
130

```

<210> 820

<211> 36

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 820

ggggaattca tgatccggga gaaatttgcc cactgc

36

<210> 821

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> PCR primer

<400> 821

gggctcgagt caggagtttg agaccagcct ggc

33

<210> 822

Lys Gln Arg Trp Gly Phe Thr Met Leu Ala Arg Leu Val Ser Asn Ser
210 215 220

<210> 826

<211> 358

<212> PRT

<213> Homo sapiens

<400> 826

Met Ser Ala Ile Glu Arg Val Ser Glu Ala Ile Val Ser Ile Arg Arg
 5 10 15

Ile Gln Thr Phe Leu Leu Leu Asp Glu Ile Ser Gln Arg Asn Arg Gln
 20 25 30

Leu Pro Ser Asp Gly Lys Lys Met Val His Val Gln Asp Phe Thr Ala
 35 40 45

Phe Trp Asp Lys Ala Ser Glu Thr Pro Thr Leu Gln Gly Leu Ser Phe
 50 55 60

Thr Val Arg Pro Gly Glu Leu Leu Ala Val Val Gly Pro Val Gly Ala
 65 70 75 80

Gly Lys Ser Ser Leu Leu Ser Ala Val Leu Gly Glu Leu Ala Pro Ser
 85 90 95

His Gly Leu Val Ser Val His Gly Arg Ile Ala Tyr Val Ser Gln Gln
 100 105 110

Pro Trp Val Phe Ser Gly Thr Leu Arg Ser Asn Ile Leu Phe Gly Lys
 115 120 125

Lys Tyr Glu Lys Glu Arg Tyr Glu Lys Val Ile Lys Ala Cys Ala Leu
 130 135 140

Lys Lys Asp Leu Gln Leu Leu Glu Asp Gly Asp Leu Thr Val Ile Gly
 145 150 155 160

Asp Arg Gly Thr Thr Leu Ser Gly Gly Gln Lys Ala Arg Val Asn Leu
 165 170 175

Ala Arg Ala Val Tyr Gln Asp Ala Asp Ile Tyr Leu Leu Asp Asp Pro
 180 185 190

Leu Ser Ala Val Asp Ala Glu Val Ser Arg His Leu Phe Glu Leu Cys
 195 200 205

Ile Cys Gln Ile Leu His Glu Lys Ile Thr Ile Leu Val Thr His Gln

Leu Ala Arg Leu Val Ser Asn Ser Leu Glu His His His His His His
 85 90 95

<210> 828
 <211> 35
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR primer

<400> 828
 cgcccatggg gatccgggag aaatttgccc actgc 35

<210> 829
 <211> 35
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR primer

<400> 829
 cgcctcgagg gagtttgaga ccagcctggc caaca 35

<210> 830
 <211> 38
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR primer

<400> 830
 gcatggacca tatgtcagcc attgagaggg tgtcagag 38

<210> 831
 <211> 34
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> PCR primer

<400> 831
 ccgctcgaga ataaggaaaa tgaagacaat ccag 34

<400> 835

```

Met His His His His His His Thr Ala Ala Ser Asp Asn Phe Gln Leu
      5                      10                      15

Ser Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala
      20                      25                      30

Ile Ala Gly Gln Ile Lys Leu Pro Thr Val His Ile Gly Pro Thr Ala
      35                      40                      45

Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val
      50                      55                      60

Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr
      65                      70                      75                      80

Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr
      85                      90                      95

Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser
      100                     105                     110

Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr
      115                     120                     125

Leu Ala Glu Gly Pro Pro Ala Glu Phe Met His Gly Pro Gln Val Leu
      130                     135                     140

Ala Arg Cys Ser Glu Cys Ala Cys Pro Ala Leu Ala Ala Thr Ser Ala
      145                     150                     155                     160

Gly Val Arg Leu Glu Gly Val Asp Arg Pro Pro Thr Leu Pro Ser Gln
      165                     170                     175

Gly Ser Gly Trp Pro Cys Ser His Ser Leu Ser Gly Cys His Leu Met
      180                     185                     190

Ala Asp Gly Ala Lys Ala Leu Gly Lys Ala Asp Gly Pro Trp Pro Tyr
      195                     200                     205

Leu Phe Val Arg Arg Thr Asp Val Pro Cys Pro Ala Ala Ser Glu Val
      210                     215                     220

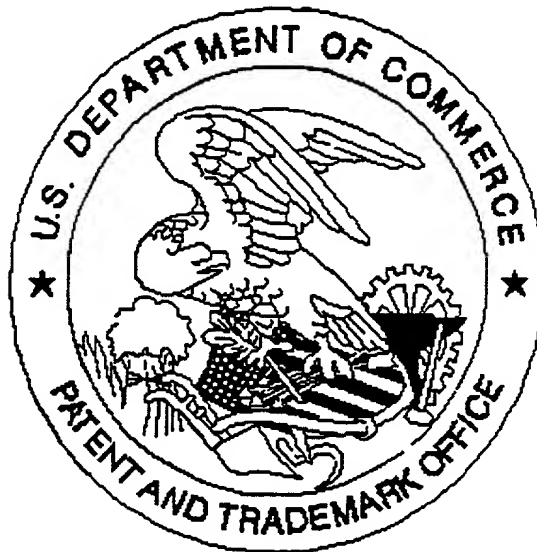
Gly Gly Cys Ala Pro Ser Ser Trp Arg Ala Leu Ala Glu Val Thr Gly
      225                     230                     235                     240

Cys Ser Leu Gly Pro Leu Gly Leu Ala Gln His Ala Gln Ala Ser Val
      245                     250                     255

Leu Leu Leu Cys Tyr Lys Trp Ser His Ile Gly Glu Thr Ser Ser His

```


United States Patent & Trademark Office
Office of Initial Patent Examination -- Scanning Division



Application deficiencies were found during scanning:

☐ Page(s) _____ of _____ were not present
for scanning. (Document title)

☐ Page(s) _____ of _____ were not present
for scanning. (Document title)

☒ Scanned copy is best available. *of drawings*

00605793-062700